

QY 61 TGCCCAAGCTAATGATATTCACAGATTGGTTAAGCAATCGTTTCTATCGAAGACCA 120
 Db 61 TGCCCAAGCTAATGATATTCACAGATTGGTTAAGCAATCGTTTCTATCGAAGACCA 120
 QY 121 TCGCTTCTCGACACAGGGGATTTAGTACCAATCGTATCTCGGAGCTTTCTCGGCAA 180
 Db 121 TCGCTTCTCGACACAGGGGATTTAGTACCAATCGTATCTCGGAGCTTTCTCGGCAA 180
 QY 181 TCTGCAAGCAATTCCTTCCAGGTGGATCAACTCTACCCCAACAGTTGATTAAGTTGAC 240
 Db 181 TCTGCAAGCAATTCCTTCCAGGTGGATCAACTCTACCCCAACAGTTGATTAAGTTGAC 240
 QY 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAGGCTCAGGAAGCTTGGTT 300
 Db 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAGGCTCAGGAAGCTTGGTT 300
 QY 301 AGCGATTTCAGTTAGCAAAAAGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360
 Db 301 AGCGATTTCAGTTAGCAAAAAGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360
 QY 361 GGTCTACATGCTAATGGGAATATGGGAATGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 420
 Db 361 GGTCTACATGCTAATGGGAATATGGGAATGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 420
 QY 421 AGACCTCAATATTAAGTTTACCTCAGTTAGCTTGGCTGGGAATGCCCTCAGGCACC 480
 Db 421 AGACCTCAATATTAAGTTTACCTCAGTTAGCTTGGCTGGGAATGCCCTCAGGCACC 480
 QY 481 AAACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGACCCGCAAGCAAGCAAGCA 540
 Db 481 AAACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGACCCGCAAGCAAGCAAGCA 540
 QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTCAACAGATGAGAAGCAAGCAAGCAAG 600
 Db 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTCAACAGATGAGAAGCAAGCAAGCAAGCA 600
 QY 601 ACCAATTTACTGATGACTACAAAGTCTCAAAATCAGCAAGTATTAACCTGCTTACATGGA 660
 Db 601 ACCAATTTACTGATGACTACAAAGTCTCAAAATCAGCAAGTATTAACCTGCTTACATGGA 660
 QY 661 TAAATTTACCTCAAGGAAGTCAATCAAGTTGAAGCAAGCAAGCAAGCAAGCAAGCAAG 720
 Db 661 TAAATTTACCTCAAGGAAGTCAATCAAGTTGAAGCAAGCAAGCAAGCAAGCAAGCAAG 720
 QY 721 AACTGGGATGATGCTACACAAATGTAGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 780
 Db 721 AACTGGGATGATGCTACACAAATGTAGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 780
 QY 781 TTAACAATACAGACGAATACGTTGCTATCCAGACGATGAATTCGAAGTCTTCTACCAT 840
 Db 781 TTAACAATACAGACGAATACGTTGCTATCCAGACGATGAATTCGAAGTCTTCTACCAT 840
 QY 841 TGTGATGCTTTACCGTAAAGTCAATTCGCCAGCTAGAGCAGCCCAATCAGTCAAGTAA 900
 Db 841 TGTGATGCTTTACCGTAAAGTCAATTCGCCAGCTAGAGCAGCCCAATCAGTCAAGTAA 900
 QY 901 TGTGATGCTTTACCGTAAAGTCAATTCGCCAGCTAGAGCAGCCCAATCAGTCAAGTAA 960
 Db 901 TGTGATGCTTTACCGTAAAGTCAATTCGCCAGCTAGAGCAGCCCAATCAGTCAAGTAA 960
 QY 961 ACCGATACAGACTATGCTTCTGCTTGGAGTACGCTTCTTGTATTAAGTGGGATAG 1020
 Db 961 ACCGATACAGACTATGCTTCTGCTTGGAGTACGCTTCTTGTATTAAGTGGGATAG 1020
 QY 1021 CGTTCACGATGAGCCCTATACTACCTTGGGCAAAATCTCTGTTTATTAAGTGGGATAG 1080
 Db 1021 CGTTCACGATGAGCCCTATACTACCTTGGGCAAAATCTCTGTTTATTAAGTGGGATAG 1080
 QY 1081 GGGCTACTTTGGCAACATCACTTTCGAATACGCTTGAACCAATCGCGAAAGCTCCGAGC 1140
 Db 1081 GGGCTACTTTGGCAACATCACTTTCGAATACGCTTGAACCAATCGCGAAAGCTCCGAGC 1140

QY 1141 CGTGAAGACTCTAAACAAGTTCGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG 1200
 Db 1141 CGTGAAGACTCTAAACAAGTTCGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG 1200
 QY 1201 AATCAGTACCAAGTATTTCACTACTCAATGCCATTTCAAGTAAACAACCAAGTACGA 1260
 Db 1201 AATCAGTACCAAGTATTTCACTACTCAATGCCATTTCAAGTAAACAACCAAGTACGA 1260
 QY 1261 CAAAAAATATGAGCAAGTAGTGAAGAGTGGCTGCTGCTTACGCTTGGCTTTGCAAGTGG 1320
 Db 1261 CAAAAAATATGAGCAAGTAGTGAAGAGTGGCTGCTGCTTACGCTTGGCTTTGCAAGTGG 1320
 QY 1321 TGAAGTCTTACTATAAACAATGTATATCCATAAAGTCGCTTTTAGTGTAGGAGTGAAGA 1380
 Db 1321 TGAAGTCTTACTATAAACAATGTATATCCATAAAGTCGCTTTTAGTGTAGGAGTGAAGA 1380
 QY 1381 AGAGTTCTTAAATGTCGGAACCTGTCGCCATGAAGAAAGCAGAGCCTATATGATGACCGA 1440
 Db 1381 AGAGTTCTTAAATGTCGGAACCTGTCGCCATGAAGAAAGCAGAGCCTATATGATGACCGA 1440
 QY 1441 CATGATGAAACAGTCTTGACTTATGGAATCGAGCAAGTGGCTATCTTGGCTGGCTCCC 1500
 Db 1441 CATGATGAAACAGTCTTGACTTATGGAATCGAGCAAGTGGCTATCTTGGCTGGCTCCC 1500
 QY 1501 TCAGGCTGGTAAACAGGAACTCTAACTATACAGACGAGGAAATTTGAAACACACATCAA 1560
 Db 1501 TCAGGCTGGTAAACAGGAACTCTAACTATACAGACGAGGAAATTTGAAACACACATCAA 1560
 QY 1561 GACCTCTCAATTTGTAGCACTGATGAACCTATTTGCTGGCTATACGCTTAAATATCAAT 1620
 Db 1561 GACCTCTCAATTTGTAGCACTGATGAACCTATTTGCTGGCTATACGCTTAAATATCAAT 1620
 QY 1621 GGCTGTATGGACAGGCTATTTAAACGCTGTACACACCTTTGAGCAATGGCTTACGGT 1680
 Db 1621 GGCTGTATGGACAGGCTATTTAAACGCTGTGTACACACCTTTGAGCAATGGCTTACGGT 1680
 QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAGA 1740
 Db 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAGA 1740
 QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGATTTTAAATTTGGTCTCG 1800
 Db 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGATTTTAAATTTGGTCTCG 1800
 QY 1801 TTCTACGTGGAACTCACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
 Db 1801 TTCTACGTGGAACTCACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
 QY 1861 ATCAGATAGTTCAACTTCAAGTCTAGCTCAACCTTCCAGCAACAAATATAGTACGAC 1920
 Db 1861 ATCAGATAGTTCAACTTCAAGTCTAGCTCAACCTTCCAGCAACAAATATAGTACGAC 1920
 QY 1921 TACCAATCCTTACAAATATACCAACAAATCAATCAACCCCTGATCAACAAATCAGAA 1980
 Db 1921 TACCAATCCTTACAAATATACCAACAAATCAATCAACCCCTGATCAACAAATCAGAA 1980
 QY 1981 TCCTCAACAGCAGCAACCA 1999
 Db 1981 TCCTCAACAGCAGCAACCA 1999

RESULT 2
 US-08-481-435-5
 ; Sequence 5, Application US/08481435
 ; Patent No. 6027906
 ; GENERAL INFORMATION:
 ; APPLICANT: Balganes, Tanjore S
 ; APPLICANT: Town, Christine
 ; TITLE OF INVENTION: No. 6027906el Polypeptides
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: White & Case
 ; STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2049 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
STRAIN: PM 1
IMMEDIATE SOURCE:
LIBRARY: PCR cloning
CLONE: PARC 0512 Soluble PBP 1A del 38
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2049
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..2046
US-08-481-435-5

Query Match 99.1%; Score 1981.4; DB 3; Length 2049;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTAGCAATAAATAAATCACTCATGCTGGTCTGAGCGCGCGTCAA 60
DB 48 TAAATCTAGCAATAAATAAATCACTCATGCTGGTCTGAGCGCGCGTCAA 107
QY 61 TGCCCAAGCTTAATGATATCCCAACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA 120
DB 108 TGCCCAAGCTTAATGATATCCCAACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA 167
QY 121 TCGCTTCTTCGACCACAGCGGGGATTGATACCATCCGTTATCCTGGAGCTTTCTTGGCGCAA 180
DB 168 TCGCTTCTTCGACCACAGCGGGGATTGATACCATCCGTTATCCTGGAGCTTTCTTGGCGCAA 227
QY 181 TCGCAAGCAATTCCTTCCCAAGGTGATCACTTCACCCCAACAGTTGATTAGTTGAC 240
DB 228 TCGCAAGCAATTCCTTCCCAAGGTGATCACTTCACCCCAACAGTTGATTAGTTGAC 287
QY 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTCTCGTAAGGCTCAGGAAGCTTGGTT 300
DB 288 TTACTTTTCAACTTCGACTTCCGACCAGACTATTCTCGTAAGGCTCAGGAAGCTTGGTT 347
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTTACTATATAATAA 360

DB 348 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTTACTATATAATAA 407
QY 361 GGTCTACATCTCTAATGGGAAGTATGGAATGCAGACAGCAGCTCAAACTACTATGTGTA 420
DB 408 GGTCTACATCTCTAATGGGAAGTATGGAATGCAGACAGCAGCTCAAACTACTATGTGTA 467
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTCTGGCTGGAAATGCTCAGGCACC 480
DB 468 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTCTGGCTGGAAATGCTCAGGCACC 527
QY 481 AAACCAATATGACCCCTATTTCACATCCAGAAGCAGCCCAAGACCGCGGAACTGGTCTT 540
DB 528 AAACCAATATGACCCCTATTTCACATCCAGAAGCAGCCCAAGACCGCGGAACTGGTCTT 587
QY 541 ATCTGAAATGAAAAATCAAGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
DB 588 ATCTGAAATGAAAAATCAAGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 647
QY 601 ACCAATTTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAAATACCTGCTTACATGGA 660
DB 648 ACCAATTTACTGATGGCTACAAAGTCTCAAAATCAGCAAGTAAATACCTGCTTACATGGA 707
QY 661 TAATTTACTCTAAGGAAGTCAATCAATCAAGTTGAAGAAAGCAAGGCTATACCTTACTCAC 720
DB 708 TAATTTACTCTAAGGAAGTCAATCAATCAAGTTGAAGAAAGCAAGGCTATACCTTACTCAC 767
QY 721 AACTGGGATGGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAGACATCTGGGGATAT 780
DB 768 AACTGGGATGGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAGACATCTGGGGATAT 827
QY 781 TTACAATACAGAGCAATACGTTGCCTATCCAGACGATGAATTTGCAAGTCGCTTCTACCAT 840
DB 828 TTACAATACAGAGCAATACGTTGCCTATCCAGACGATGAATTTGCAAGTCGCTTCTACCAT 887
QY 841 TGTGTGATGTTTCTAAGCGTAAAGTCAATGCCCAGCTAGGAGCAGGCATCAGTCAAGTAA 900
DB 888 TGTGTGATGTTTCTAAGCGTAAAGTCAATGCCCAGCTAGGAGCAGGCATCAGTCAAGTAA 947
QY 901 TGTGTGATGTTTCTAAGCGTAAAGTCAATGCCCAGCTAGGAGCAGGCATCAGTCAAGTAA 960
DB 948 TGTGTGATGTTTCTAAGCGTAAAGTCAATGCCCAGCTAGGAGCAGGCATCAGTCAAGTAA 1007
QY 961 ACCGATCACAGACTATGCTCTGCTTGGAGTACGGTGTCTAGGATTTCAACTGCTACTAT 1020
DB 1008 ACCGATCACAGACTATGCTCTGCTTGGAGTACGGTGTCTAGGATTTCAACTGCTACTAT 1067
QY 1021 CGTTACGATGAGCCCTTAACTACCCCTGGGCAAAATACCTCTGTTTATTAACCTGGGATAG 1080
DB 1068 CGTTACGATGAGCCCTTAACTACCCCTGGGCAAAATACCCCTGTTTATTAACCTGGGATAG 1127
QY 1081 GGCTACTTTGGCAACATACCTTGCATACGCCCTGCAACATCGCGAAGCTGCCAGC 1140
DB 1128 GGCTACTTTGGCAACATACCTTGCATACGCCCTGCAACATCGCGAAGCTGCCAGC 1187
QY 1141 CGTGGAAACTCTAAACAAGGTCGAGCTCAACCGCGCAAGACTTTTCCCTAAATGGTCTAGG 1200
DB 1188 CGTGGAAACTCTAAACAAGGTCGAGCTCAACCGCGCAAGACTTTTCCCTAAATGGTCTAG 1247
QY 1201 AATCGACTACCCCAAGTATTCACACTACTCAAAATGCCATTTCAGTAACACACCAAGTAA 1260
DB 1248 AATCGACTACCCCAAGTATTCACACTACTCAAAATGCCATTTCAGTAACACACCAAGTAA 1307
QY 1261 CAAAAATATGGAGCAAGTAGTGAAGAGTGGCTGCTGCTAGCTGCCCTTTGCAATGG 1320
DB 1308 CAAAAATATGGAGCAAGTAGTGAAGAGTGGCTGCTGCTAGCTGCCCTTTGCAATGG 1367
QY 1321 TGAACCTTACTATAAACCAATGATATCCATAAAGTCTCTTTAGTGTAGGAGTGAAGA 1380
DB 1368 TGAACCTTACTATAAACCAATGATATCCATAAAGTCTCTTTAGTGTAGGAGTGAAGA 1427
QY 1381 AGAGTTCTCTAATGTCGGAAGTCTGTCATGAAGGAAAGCAGCCCTTATGATGACCGA 1440

Db 1426 AGAGTTCTCTTCTGCGAACTGTCGCCATGAAAGAAAGCAGACGCTATATGATGACCGA 1487
 QY 1441 CATGATGAAGACAGTCTTACATATGGAAGTGGACCAATGCTATCTTGGCTGCC 1500
 Db 1488 CATGATGAAGACAGTCTTACATATGGAAGTGGACCAATGCTATCTTGGCTGCC 1547
 QY 1501 TCAGGCTGTAAGACAGGAACTCTTAACATATACAGACGAGGAAATGAAACCAATCAA 1560
 Db 1548 TCAGGCTGTAAGACAGGAACTCTTAACATATACAGACGAGGAAATGAAACCAATCAA 1607
 QY 1561 GACCTCTCAATTTGTAGACACCTGATGAACTATTTGCTGGCTATACCGTAAATTTCAAT 1620
 Db 1608 GACCTCTCAATTTGTAGACACCTGATGAACTATTTGCTGGCTATACCGTAAATTTCAAT 1667
 QY 1621 GCGTGTATGACAGGCTATTTAAACGCTCTGACACCACTGTAGGCAATGGCTTACGCT 1680
 Db 1668 GCGTGTATGACAGGCTATTTAAACGCTCTGACACCACTGTAGGCAATGGCTTACGCT 1727
 QY 1681 CGCTGCAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAGA 1740
 Db 1728 CGCTGCAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAGA 1787
 QY 1741 TTGGAATATACAGAGGCTCTACAGAAATGGAGAAATTCGTAATTTAAAAATGGTCTCG 1800
 Db 1788 TTGGAATATACAGAGGCTCTACAGAAATGGAGAAATTCGTAATTTAAAAATGGTCTCG 1847
 QY 1801 TTCTAGTGAACCTACCTGCTCCACAAACCCCATCACTGAAAGTTCAAGCTCATC 1860
 Db 1848 TTCTAGTGAAGCTACCTGCTCCACAAACCCCATCACTGAAAGTTCAAGCTCATC 1907
 QY 1861 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCCCAAGCACAATAATAGTACGAC 1920
 Db 1908 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCCCAAGCACAATAATAGTACGAC 1967
 QY 1921 TACCAATCTACATAATACGAGCAATCAATCAACCACTCCCAAGCACAATAATAGTACGAC 1980
 Db 1968 TACCAATCTACATAATACGAGCAATCAATCAACCACTCCCAAGCACAATAATAGTACGAC 2027
 QY 1981 TCCTCAACCAAGCACAACCA 1999
 Db 2028 TCCTCAACCAAGCACAACCA 2046

RESULT 3
 US-08-245-511-3
 ; Sequence 3, Application US/08245511
 ; Patent No. 5928900
 ; GENERAL INFORMATION:
 ; APPLICANT: Masure, H Robert
 ; APPLICANT: Pearce, Barbara J
 ; APPLICANT: Tuomanen, Elaine
 ; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
 ; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
 ; NUMBER OF SEQUENCES: 58
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klauber & Jackson
 ; STREET: 411 Hackensack Avenue
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/245,511
 ; FILING DATE: 18-MAY-1994
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/116,541
 ; FILING DATE: 01-SEP-1994

ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 600-1-069 CIP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201 487-5800
 ; TELEFAX: 201 343-1684
 ; TELEX: 133521
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 960 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Streptococcus pneumoniae
 ; STRAIN: R6
 ; IMMEDIATE SOURCE:
 ; CLONE: SPRU42
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..960
 ; US-08-245-511-3

Query Match 47.2%; Score 944.2; DB 2; Length 960;
 Best Local Similarity 99.7%; Pred. No. 1.9e-278;
 Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TAAATCTAGCAATAAATAAATCAACTCATCTGCTGACTTGGGTTCTGAAGCGCGCTCAA 60
 Db 12 TAAATCTAGCAATAAATAAATCAACTCATCTGCTGACTTGGGTTCTGAAGCGCGCTCAA 71
 QY 61 TGCCCAAGCTAATGATATTTCCACACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
 Db 72 TGCCCAAGCTAATGATATTTCCACACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 131
 QY 121 TCGCTTCTTCGACCACAGGGGATTTGATACCATCCCTATCTCTGGGAGCTTCTTGGCGCA 180
 Db 132 TCGCTTCTTCGACCACAGGGGATTTGATACCATCCCTATCTCTGGGAGCTTCTTGGCGCA 191
 QY 181 TCTGCAAAAGCAATTCCTCCCAAGTGGATCACTCACTCAACCAACAGTTGATTAAGTTGAC 240
 Db 192 TCTGCAAAAGCAATTCCTCCCAAGTGGATCACTCACTCAACCAACAGTTGATTAAGTTGAC 251
 QY 241 TTACTTTTCAACTTCGACTTCGACATTCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
 Db 252 TTACTTTTCAACTTCGACTTCGACATTCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 311
 QY 301 AGGATTTAGTGAACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360
 Db 312 AGGATTTAGTGAACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 371
 QY 361 GGTCTACATGCTTAATGGGAACCTATGGAATGCAAGCAGAGCTCAAAAGCTACTATGTTAA 420
 Db 372 GGTCTACATGCTTAATGGGAACCTATGGAATGCAAGCAGAGCTCAAAAGCTACTATGTTAA 431
 QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGATGGAATGCTCAGGCACC 480
 Db 432 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGATGGAATGCTCAGGCACC 491
 QY 481 AAACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGACCCGCGGAACTTGGTCTT 540
 Db 492 AAACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGACCCGCGGAACTTGGTCTT 551
 QY 541 ATCTGAATGAAATAAATCAAGGCTTACATCTCTGCTGAACAGTATGAGAAAGCAGCTCAATAC 600
 Db 552 ATCTGAATGAAATAAATCAAGGCTTACATCTCTGCTGAACAGTATGAGAAAGCAGCTCAATAC 611
 QY 601 ACCAATTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTTACCCTGCTTACATGGA 660

Db 612 ACCAATTACTGATGGGCTCAAAAGTCTCAATCAGCAAGTAATTACCCCTGCTTACATGGA 671
QY 661 TAATTACCTCAAGGAAGTATCAATCAAGTTGAAGAGAAACAGGCTATTAACCTACTCAC 720
Db 672 TAATTACCTCAAGGAAGTATCAATCAAGTTGAAGAGAAACAGGCTATTAACCTACTCAC 731
QY 721 AACTGGGATGGATGCTACACAAATAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
Db 732 AACTGGGATGGATGCTACACAAATAGACCAAGAGCTCAAAAACATCTGTGGGATAT 791
QY 781 TTACAATACAGCAAGTAAGTTGCCCTATCCAGACGATGAATTCGAAGTGGCTTCTACCAT 840
Db 792 TTACAATACAGCAAGTAAGTTGCCCTATCCAGACGATGAATTCGAAGTGGCTTCTACCAT 851
QY 841 TGTGTATGTTCTTAACGGTTAAAGTCAATGCCAGCTAGGAGCAGCGCATCAGTCAAGTAA 900
Db 852 TGTGTATGTTCTTAACGGTTAAAGTCAATGCCAGCTAGGAGCAGCGCATCAGTCAAGTAA 911
QY 901 TGTTCCTTCGGGAATTAACCAAGCAGTAGAACAACACCGGACTGGGGA 949
Db 912 TGTTCCTTCGGGAATTAACCAAGCAGTAGAACAACACCGGACTGGGGA 960

RESULT 4

US-08-600-993A-3
; Sequence 3, Application US/08600993A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 960 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: R6
; IMMEDIATE SOURCE:
; CLONE: SPRU42
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..960
; US-08-600-993A-3

Query Match 47.2%; Score 944.2; DB 2; Length 960;
Best Local Similarity 99.7%; Pred. No. 1.9e-278;
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TAAATCTAGGACAATAAAATCAACTCATGCTGACTGGGTCTGACGCGCGGTCAA 60
Db 12 TAAATCTAGGACAATAAAATCAACTCATGCTGACTGGGTCTGACGCGCGGTCAA 71
QY 61 TGCCCAAGCTTAATGATATCCACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA 120
Db 72 TGCCCAAGCTTAATGATATCCACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA 131
QY 121 TCGCTTCTTCGACCACAGGGGGATTGATACATCCGCTATCTCGGGAGCTTCTTGGCGCA 180
Db 132 TCGCTTCTTCGACCACAGGGGGATTGATACATCCGCTATCTCGGGAGCTTCTTGGCGCA 191
QY 181 TGTGCAAGCAATTCCTCCCAAGGTGGATCACTCAACCAACAGTTGATTAAGTTGAC 240
Db 192 TGTGCAAGCAATTCCTCCCAAGGTGGATCACTCAACCAACAGTTGATTAAGTTGAC 251
QY 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 252 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 311
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
Db 312 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 371
QY 361 GGTCTACATGCTTAATGGGAACCTATGGAATGCAGACAGCAGCTCAAAACCTACTATGATA 420
Db 372 GGTCTACATGCTTAATGGGAACCTATGGAATGCAGACAGCAGCTCAAAACCTACTATGATA 431
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGTGGTGGATGCTCAGGCACC 480
Db 432 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGTGGTGGATGCTCAGGCACC 491
QY 481 AAACCAATATGACCCCTATTTCACATCCAGAACGAGCCCAAGACCCGCGAAACCTTGGTCTT 540
Db 492 AAACCAATATGACCCCTATTTCACATCCAGAACGAGCCCAAGACCCGCGAAACCTTGGTCTT 551
QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGATATGAGAAAGCAGCTCAATAC 600
Db 552 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGATATGAGAAAGCAGCTCAATAC 611
QY 601 ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAAGTAAATTCCTGCTTACATGGA 660
Db 612 ACCAATTACTGATGGGCTACAAAGTCTCAATCAGCAAGTAAATTCCTGCTTACATGGA 671
QY 661 TAATTACCTCAAGGAAGTATCAATCAAGTTGAAGAGAAACAGGCTATAAACCTACTCAC 720
Db 672 TAATTACCTCAAGGAAGTATCAATCAAGTTGAAGAGAAACAGGCTATAAACCTACTCAC 731
QY 721 AACTGGGATGGATGCTACACAAATAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
Db 732 AACTGGGATGGATGCTACACAAATAGACCAAGAGCTCAAAAACATCTGTGGGATAT 791
QY 781 TTACAATACAGCAAGTAAGTTGCCCTATCCAGACGATGAATTCGAAGTGGCTTCTACCAT 840
Db 792 TTACAATACAGCAAGTAAGTTGCCCTATCCAGACGATGAATTCGAAGTGGCTTCTACCAT 851


```

Db 200 GCTTCTACAGCATCAGCGCTTGACCCGGTGGGATCTTCGTCGAGCAAGCGTGGCGC 259
QY 183 TGCAGAGCAATCCCTCCCAAGTGGATCAACTCTACCCCAAGCAAGTGGATTAAGTTGACTT 242
Db 260 TGTCTCCGGTCACGGCTCAAGGGGCAAGTACCAATACCAGCAGCTGGCGAGAACT 319
QY 243 ACTTTTCAACTTCGACTTCGCCACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGTTAG 302
Db 320 TCTTCTCAGTCC-----AGAACGACGCTGATGCGTAAGATTAAAGGAAGTCTTCTCTCG 373
QY 303 CGATTTCAGTTAGCAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAAGG 362
Db 374 CGATTTCGCTTGAACCTGCTGACGAAAAGCAGAGATCCTCGAGCTTTATCTGAACAAGA 433
QY 363 TCTACATGCTTAATGGAACTATGGAATGCAGACAGCAGCTCAAAACTACTATATGTAAG 422
Db 434 TTTACCTTGGTTACCGCGCTATGGTGTGCTGCGGCAAGTCTATTTCCGGAAGAA 493
QY 423 ACCTCAATATTAAGTTTACCTCAGTTAGCCTTGCTGCTGGAATGSCCTCAGGCACCAA 482
Db 494 CGGTCCACCAACTGACGCTGAACGAATGCGGTGATAGCGGCTGCCGAAAGCGCCTT 553
QY 483 ACCAATATACCCCTATTACATCCAGAGCAGCCCAAGCCGCGCAAGCTTGCTCTTAT 542
Db 554 CCACCTTCAACCCGCTCTACTCGATGGATCGTCCGCTCGCGGCGTAAACGTCGTCTGT 613
QY 543 CTGAATGAAAATCAAGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATACAC 602
Db 614 CCGGATGCTGATGAAGGTTATATACCCACAAAGTTCGATCAGACAGCAGCTGAGG 673
QY 603 CAATTACTGAT 613
Db 674 CGATTAACGCT 684

```

RESULT 8

```

US-08-771-716-1
; Sequence 1, Application US/08771716
; Patent No. 5922540
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Hoskins, JoAnn
; APPLICANT: Jaskunas, S.Richard
; APPLICANT: Skatrud, Paul L.
; TITLE OF INVENTION: Monofunctional Glycosyltransferase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/771,716
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: x-11067
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..807
US-08-771-716-1

Query Match      3.1%; Score 62.2; DB 2; Length 807;
Best Local Similarity 49.9%; Pred. No. 3.4e-09;
Matches 191; Conservative 0; Mismatches 183; Indels 9; Gaps 1;

QY 67 ACCTAATGATATCCACAGATTGGTTAAGCAATCGTTTCTATCGAAGACCATCGCTT 126
Db 252 AGCTGATAACATGCCAGAGTATGTTAAAGTGCTTTATTTCAATGAAGATGAACGATT 311
QY 127 CTTGACACAGAGGGGATGATACCATCGGTATCTCTGGGAGCTTTCTTGGCAATCTGCA 186
Db 312 CTACATCATCATGATTGATTTGAAGGTACAAGTATTTTCAACGATTAG 371
QY 187 AAGCAATTCCTCCCAAGTGGATCAACTCTCACCAACAGTTGATTAAGTTGACTTACTT 246
Db 372 CGACAGAGATGTGCAAGTGGTAGTACCATTACACAAAGTTGTCAAAAATTTATTTT 431
QY 247 TTCACTTCGACTTCGACGACGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTTAGCAT 306
Db 432 TGATAAT-----GATCGTTTACTTACTAGAAAGTAAAGAAATTAATTTAGCTCA 482
QY 307 TCAGTTAGAACAAAAAGCAACCAAGAAATCTTGACCTACTATATAAATAGGTCTA 366
Db 483 TCGAGTTGAAAAACAATATAATAAGACGAAATTTTAAGCTTTTATTAATAATATTTA 542
QY 367 CATGCTCTAATGGAACTATGGAATGCAGACAGCAGCTCAAAACTACTATGTTAAAGACCT 426
Db 543 CTTTGGGGATAATCAATATACGCTTGAGGCGCAGCAAAACCATTTCTTGGAAACACCGT 602
QY 427 CAATAATTTAAGTTTACCTCAGT 449
Db 603 GAATAAAAAATAGTACCAACATGT 625

```

RESULT 9

```

US-08-771-716-3
; Sequence 3, Application US/08771716
; Patent No. 5922540
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Hoskins, JoAnn
; APPLICANT: Jaskunas, S.Richard
; APPLICANT: Skatrud, Paul L.
; TITLE OF INVENTION: Monofunctional Glycosyltransferase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/771,716
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

```



```

367 CATGCTCTAATGGGAACTATGGAATGCAGACGAGCAGCTCAAAAGACTTACTATGTAACA
    || || || || || || || || || || || || || || || || || || || || ||
Db 543 CTTTGGGGATAATCAATATACGCTTGAGGGCGCAGCAACCATTTACTTTGGAACA
    || || || || || || || || || || || || || || || || || || || || ||
Qy 427 CAATAATTTTAAGTTTACCTCAGT 449
    ||||| ||| | | ||
Db 603 GAATAAAAAATAGTACAACAATGT 625
    ||||| ||| | | ||

RESULT 11
US-09-057-720A-3
; Sequence 3, Application US/09057720A
; Patent No. 6143868
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Hoskins, JoAnn
; APPLICANT: Jaskunas, S.Richard
; APPLICANT: Skatrud, Paul L.
; TITLE OF INVENTION: Monofunctional Glycosyltransferase Gene
; TITLE OF INVENTION: of Staphylococcus Aureus
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company

```



```

: STREET: Lilly Corporate Center
: CITY: Indianapolis
: STATE: Indiana
: COUNTRY: U.S.
: ZIP: 46285
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/09/057,720A
:

```

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11067
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-09-057-720A-3

```

Query Match

Best Local Similarity 33.9%; Pred. No. 3.4e-09;
Matches 130; Conservative 61; Mismatches 183; Indels 9; Gaps 1;

[illegible]

RESULT 12

RESULT 12
US-08-731-716-1
; Sequence 1, Application US/08731716
; Patent No. 5789202
; GENERAL INFORMATION:
; APPLICANT: Hoskins, JoAnn
; APPLICANT: Jaskunas, S. Richard
; APPLICANT: Rockney, Pamela K.

```

> APPLICANT: Zhao, Genshi
> APPLICANT: Rosteck, Paul R. Jr.
> APPLICANT: No. 5789202ris, Franklin H.
> TITLE OF INVENTION: Penicillin Binding Protein From
> TITLE OF INVENTION: Streptococcus Pneumoniae
> NUMBER OF SEQUENCES: 3
> CORRESPONDENCE ADDRESS:
> ADDRESSEE: Eli Lilly and Company
> STREET: Lilly Corporate Center
> CITY: Indianapolis
> STATE: Indiana
> COUNTRY: U.S.
> ZIP: 46285
> COMPUTER READABLE FORM:
> MEDIUM TYPE: Floppy disk
> COMPUTER: IBM PC compatible
> OPERATING SYSTEM: PC-DOS/MS-DOS
> SOFTWARE: PatentIn Release #1.0, Version #1.30
> CURRENT APPLICATION DATA:
> APPLICATION NUMBER: US/08/731.716
> FILING DATE:
> CLASSIFICATION: 435
> ATTORNEY/AGENT INFORMATION:
> NAME: Webster, Thomas D.
> REGISTRATION NUMBER: 39,872
> REFERENCE/DOCKET NUMBER: X-10,887
> TELECOMMUNICATION INFORMATION:
> TELEPHONE: 317-276-3334
> INFORMATION FOR SEQ ID NO: 1:
> SEQUENCE CHARACTERISTICS:
> LENGTH: 2193 base pairs
> TYPE: nucleic acid
> STRANDEDNESS: single
> TOPOLOGY: linear
> MOLECULE TYPE: DNA (genomic)
> HYPOTHEetical: NO
> ANTI-SENSE: NO
> FEATURE:
> NAME/KEY: CDS
> LOCATION: 1..2193
>
US-08-731-716-1

```

Query Match
2.8%; Score 56.4; DB 1; Length 2193;

Query Match	2.0%	Score: 44.7	Length: 2157
Best Local Similarity	44.6%	Pred. No. 3.4e-07;	
Matches 381: Conservative	0;	Mismatches 446;	Indels 27; Gaps 3;

[illegible]


```

Db 625 TATTTTGGAAATGGTGTGTGGGGGTAGAAAGATGCGAGTAAAGAAATACATTGGAGTTCT 684
QY 425 CTCAATAATTTAAGTTTACCTCAGTTAGCTTGTCTGGCTGGAAATGCTCAGGCACCAAAAC 484
Db 685 GCATCAGAGTGTGCTGATCAAGCTGCGACTCTGGCAGGATGCTCAAGGGCGGGA 744
QY 485 CAATATGACCCCTATTTCATCCAGAACGAGCCGAGCCGAGAACTTGTGCTTTATCT 544
Db 745 CTGTATAATCCCTTGAATTCCTGAGAAATCTTACTAATCGGCGGATACTGCTTTCAG 804
QY 545 GAAATGAAAAATCAGGCTTACATCTCTGCTGACAGTATGAGAAAGCAGCTCAATACCA 604
Db 805 ATATGTTGGACGAGATATATTGATAAAACCCAGAAACCGAAGCTGCTGAAGTTGAT 864
QY 605 ATTACTGATGGACTCAAAAGTCTCAAAATCAGCAAGTAAATACCCCTGTTTACATGGATA 664
Db 865 ATGACTTCGCAATTGCACGATAGTATGAAGGAAAAATCTCAGATTACCGTTACCCCTCT 924
QY 665 TACCTCAAGGAAGTCAATCAAGTTG-----AAGAAGAAACAGGCTATACCTA 715
Db 925 TATTTTGTATGCGGTGGTTAATGAAGCTGTTTCCAAATATTAATCTACAGAGGAAGATT 984
QY 716 CTCACAACCTGGGATGATCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGG 775
Db 985 GTCAATAATGGCTACCGCATTTACACAGAGCTGGACCAAAACTACCAACAAATATGCGAG 1044
QY 776 GATATTTACAATACAGACGAATACGTTGCTTATCCAGACGATGAATTCGAAGTCGCTTCT 835
Db 1045 ATTGTTTATGAACAACATCGCTATTTCGAGGGCAGAGGATGGAACGTTTGTCTCAATCA 1104
QY 835 ACCATTGTTGATGT 849
Db 1105 GGAAGTGTAGCTCT 1118

```

RESULT 13

```

US-08-731-716-3
; Sequence 3, Application US/08731716
; Patent No. 5789202
; GENERAL INFORMATION:
; APPLICANT: Hoskins, JoAnn
; APPLICANT: Jaskunas, S. Richard
; APPLICANT: Rocky, Pamela K.
; APPLICANT: Zhao, Genshi
; APPLICANT: Rostek, Paul R. Jr.
; APPLICANT: No. 5789202is, Franklin H.
; TITLE OF INVENTION: Penicillin Binding Protein From
; TITLE OF INVENTION: Streptococcus Pneumoniae
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,716
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-10,887
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 2193 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-731-716-3

Query Match          2.8%; Score 56.4; DB 1; Length 2193;
Best Local Similarity 31.6%; Pred. No. 3.4e-07;
Matches 270; Conservative 111; Mismatches 446; Indels 27; Gaps 3;

QY 5 ATCTAGCAACAATAAATAAATCAATCTGCTGACTTGGTCTTGGTTCGAAACCGCGCTCAATGCC 64
Db 283 AUUUUUGACCGGAGAGAAAAAGAGCGGUGGCUUGUGUCUAAAAGGAAACCUAUGUU 342
QY 65 CAAGCTAATGATATTCCCAACAGATTGTTGGTTAAGGCAATCGTTTCTATCGAAGACCAATCG 124
Db 343 GAGCUGACUGACAUCAGUAAAAACUUGCAGAAUGUGUUAUUGCGACAGAAACCCGUCU 402
QY 125 TTCTTCGACCAACAGGGGGGATTTGATACCATCCGTATCTCTGGAGCTTCTTTCGCAATCTG 184
Db 403 UUCUUAUAAAAUAGACGGGAUUUACUUAUGGCCGUUUC-----UUCUUGGCUAUGUC 453
QY 185 CAAAGCAATTCCTCCTCAAGTGTGATCAACTCTCACCAACAGTTGATTAAAGTTGACTTAC 244
Db 454 ACUGUGGAGGCUUACAGUGGUGGUCUACCAUUAUCCCAACAGCUGGCUAAAAACGCCUUA 513
QY 245 TTTTCAACTTCGACTTCCAGCAGACTATTCTCTGTAAGGCTCAGGAAGCTTGGTTAGCG 304
Db 514 UUAU-----CGCAGGAUCAACUGUUGAGAGAAAAAGCGAAAGAAUUUUUCCUUGCC 564
QY 305 ATTCAGTTAAGACAAAAACCAACCAAGCAAGAAATCTTCACCTACTATATAATAAGTGC 364
Db 565 UUGAAAUUUAAGCAAAAAAUUAUAGUAAAGGAGCAAAUUAUCCAUUACCUUACCAACGCU 624
QY 365 TACATGCTCTAATGGGAACCTATGGAATGCAGACAGCAGCTCAAAACCTACTATGTAAGAC 424
Db 625 UAUUUUGGAUUGGUGUGUGGUGGUGUAGAAGUAGAGUAGAAUAUACUUGGAGUUUCU 684
QY 425 CTCATAAATTTAAGTTTACCTCAGTTAGCTTGTCTGGCTGGAAATGCTCAGGCACCAAAAC 484
Db 685 GCAUCAAGAGAGUGAGUGUGAUAAGCUGGACUCUGGCGAGGAGUAGUACAAGGGCGGAA 744
QY 485 CAATATGACCCCTATTTCATCCAGAACGAGCCCAAGACCGCGGAACTTGGTCTTATCT 544
Db 745 CUGUAUAUCCCUUGAAUUCUGUAGAAUUAUUAUUCGCGCGGAGUACUGUCUUGCAG 804
QY 545 GAAATGAAAAATCAAGGCTTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATACCA 604
Db 805 AAUUGGUUGCAGCAGCAUUAUUAUUAUAAACCAAGAAACCGAGCUGCUGAUGUAGU 864
QY 605 ATTACTGATGGACTCAAAAGTCTCAAAATCAGCAAGTAAATTCCTCTGTTTACATGGATA 664
Db 865 AUGACUUCGCAUUGCAGCAUUAAGUAGAAAGAAAAUUCAGAUUACCUUACCCUCU 924
QY 665 TACCTCAAGGAAGTCAATCAAGTTG-----AAGAAGAAACAGGCTATACCTA 715
Db 925 UAUUUUGAGCGGUGUUAUUAAGGCGUUCUCAAAGUUAUUAUACAGAGGAAGAGUU 984
QY 716 CTCACAACCTGGGATGATCTCTACACAATGTAGACCAAGAGCTCAAAACATCTGTGG 775
Db 985 GUCAUAUUGGCUUACCGCAUUAUACAGAGCUGGCGGACCAAAACUACCAACAAUUAUGCAG 1044
QY 776 GATATTTACAATACAGACGAATACGTTGCTTATCCAGACGATGAATTCGAAGTCGCTTCT 835
Db 1045 AUUGUUUAUAGAAACACACUUGCUUAUUCGAGGCGAGAGUAGAACGUUUCUUAUCA 1104
QY 836 ACCATTGTTGATGT 849
Db 1105 GGAAGUGUAGCUCU 1118

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 12:37:03 ; Search time 3441.19 seconds
(without alignments)
12567.793 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATCTAGCACAATAAAA.....ATCCTCAACGACACAACCA 1999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 21979536 seqs, 10817449327 residues 43959072
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1:	/cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:*
2:	/cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3:	/cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4:	/cgn2_6/ptodata/2/pna/US08_COMB.seq:*
5:	/cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6:	/cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7:	/cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8:	/cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9:	/cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10:	/cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11:	/cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12:	/cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13:	/cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14:	/cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15:	/cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16:	/cgn2_6/ptodata/2/pna/US092_COMB.seq:*
17:	/cgn2_6/ptodata/2/pna/US093_COMB.seq:*
18:	/cgn2_6/ptodata/2/pna/US094_COMB.seq:*
19:	/cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
20:	/cgn2_6/ptodata/2/pna/US095B_COMB.seq:*
21:	/cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
22:	/cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
23:	/cgn2_6/ptodata/2/pna/US096A_COMB.seq:*
24:	/cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
25:	/cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
26:	/cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
27:	/cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
28:	/cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
29:	/cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
30:	/cgn2_6/ptodata/2/pna/US098A_COMB.seq:*
31:	/cgn2_6/ptodata/2/pna/US098B_COMB.seq:*
32:	/cgn2_6/ptodata/2/pna/US098C_COMB.seq:*
33:	/cgn2_6/ptodata/2/pna/US099A_COMB.seq:*
34:	/cgn2_6/ptodata/2/pna/US099B_COMB.seq:*
35:	/cgn2_6/ptodata/2/pna/US099C_COMB.seq:*
36:	/cgn2_6/ptodata/2/pna/US100_COMB.seq:*
37:	/cgn2_6/ptodata/2/pna/US101_COMB.seq:*
38:	/cgn2_6/ptodata/2/pna/US6000_COMB.seq:*
39:	/cgn2_6/ptodata/2/pna/US6001_COMB.seq:*
40:	/cgn2_6/ptodata/2/pna/US6002_COMB.seq:*
41:	/cgn2_6/ptodata/2/pna/US6003_COMB.seq:*
42:	/cgn2_6/ptodata/2/pna/US6004_COMB.seq:*
43:	/cgn2_6/ptodata/2/pna/US6005_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1999	100.0	1999	20	US-09-536-784-1	Sequence 1, Appli
2	1999	100.0	1999	30	US-09-765-271-1	Sequence 1, Appli
3	1999	100.0	1999	30	US-09-765-272-1	Sequence 1, Appli
4	1999	100.0	10383	41	US-60-029-960-81	Sequence 81, Appl
c	1999	100.0	10711	13	US-08-961-527-145	Sequence 145, App
6	1987.8	99.4	2160	22	US-09-583-110-1312	Sequence 1102, Ap
7	1987.8	99.4	2166	15	US-09-107-433-1102	Sequence 9325, Ap
8	1981.4	99.1	2160	1	PCT-US02-03987-9325	Sequence 9325, Ap
9	1981.4	99.1	2160	31	US-09-815-242-9325	Sequence 9325, Ap
10	1981.4	99.1	2160	37	US-10-072-851-9325	Sequence 9325, Ap
c	1981.4	99.1	10333	45	US-60-061-998-596	Sequence 596, App
12	1631.2	81.6	9845	45	US-60-068-175-596	Sequence 596, App
13	944.2	47.2	960	5	US-08-116-541-3	Sequence 3, Appli
14	808.6	40.5	2172	1	PCT-US97-14436-139	Sequence 139, App
15	808.6	40.5	2172	13	US-08-911-503-139	Sequence 139, App
16	808.6	40.5	2172	13	US-08-911-503A-139	Sequence 6631, Ap
17	574.6	28.7	2337	1	PCT-US02-03987-6631	Sequence 6631, Ap
18	574.6	28.7	2337	31	US-09-815-242-6631	Sequence 1534, Ap
19	574.6	28.7	2337	37	US-10-072-851-6631	Sequence 1534, Ap
20	573	28.7	2370	15	US-09-134-000-1534	Sequence 217, App
21	559	28.0	8395	14	US-09-070-927-217	Sequence 217, App
22	559	28.0	8395	14	US-09-070-927A-217	Sequence 2013, Ap
23	553	27.7	2472	15	US-09-107-532A-2013	Sequence 495, App
24	553	27.7	2472	15	US-09-107-532A-2013	Sequence 495, App
25	496.2	24.8	1386	37	US-10-091-007-91	Sequence 91, Appl
26	410	20.5	2470	45	US-60-068-186-495	Sequence 495, App
27	410	20.5	2474	44	US-60-050-444-495	Sequence 495, App
c	352.2	17.6	3478	45	US-60-068-217-710	Sequence 710, App
29	348.6	17.4	3279	24	US-09-634-238-44	Sequence 14, Appl
30	348.6	17.4	7818	24	US-09-634-238-14	Sequence 108, App
31	273	13.7	933	43	US-60-045-649-108	

Qy	1201	AATCGACTACCCCAAGTATTTCAC	TACTACTCAATGCCATTTCAAGT	TACACACCAACCGAATCAGA	1261	
Db	1201	AATCGACTACCCCAAGTATTTCAC	TACTACTCAATGCCATTTCAAGT	TACACACCAACCGAATCAGA	1261	
Qy	1261	CAAAAATATGGAGCAAGTAGTGA	AAAGATGGCTGCTTACGCTGCCT	TTCGCAATGG	1320	
Db	1261	CAAAAATATGGAGCAAGTAGTGA	AAAGATGGCTGCTTACGCTGCCT	TTCGCAATGG	1320	
Qy	1321	TGGAAC	TACTATAAACCAATGTATATC	ATAAAGTCGCTTTAGTATGGGAGTGA	1380	
Db	1321	TGGAAC	TACTATAAACCAATGTATATC	ATAAAGTCGCTTTAGTATGGGAGTGA	1380	
Qy	1381	AGAGT	TCTCTAATGTCGGAAC	TCGTGCCATGAAGAAACGACGCC	TATATGATGACCGA	1440
Db	1381	AGAGT	TCTCTAATGTCGGAAC	TCGTGCCATGAAGAAACGACGCC	TATATGATGACCGA	1440
Qy	1441	CATGATGAAAACAGTCTTGAC	TATATGGAAC	TGGAGCAATGCC	TATCTGCTTGGCTCCC	1500
Db	1441	CATGATGAAAACAGTCTTGAC	TATATGGAAC	TGGAGCAATGCC	TATCTGCTTGGCTCCC	1500
Qy	1501	TCAGGCTGGTAAAACAGGAAC	CTTAAC	TATACAGACGAGGAAAT	TGAAACACCATCAA	1560
Db	1501	TCAGGCTGGTAAAACAGGAAC	CTTAAC	TATACAGACGAGGAAAT	TGAAACACCATCAA	1560
Qy	1561	GACCTCTCAATTTGTAGCACT	TGATGCACTATT	TGCTGSC	TATACGCGTAAATATTC	1620
Db	1561	GACCTCTCAATTTGTAGCACT	TGATGCACTATT	TGCTGSC	TATACGCGTAAATATTC	1620
Qy	1621	GGCTGTATGGACAGGCTATT	TCTAACCGTCTGACAC	CACCTTGTAGGCAAT	GGCCTTACGGT	1680
Db	1621	GGCTGTATGGACAGGCTATT	TCTAACCGTCTGACAC	CACCTTGTAGGCAAT	GGCCTTACGGT	1680
Qy	1681	CGCTGCCAAAGTTTACCGTCT	TATGATGACCT	TACCTGTCTGAAGGAGCAAT	TCGAGGA	1740
Db	1681	CGCTGCCAAAGTTTACCGTCT	TATGATGACCT	TACCTGTCTGAAGGAGCAAT	TCGAGGA	1740
Qy	1741	TTGGAATATACCAAGAGGGCT	CTACAGAAATGGAGAA	TTCCG	TATTTAAAATGGTGCTCG	1800
Db	1741	TTGGAATATACCAAGAGGGCT	CTACAGAAATGGAGAA	TTCCG	TATTTAAAATGGTGCTCG	1800
Qy	1801	TTCTACGTGGAACTACCTGCT	CTCCACAAACACCC	CCCATCAACTGAAAG	TTCAAGCTCATC	1860
Db	1801	TTCTACGTGGAACTACCTGCT	CTCCACAAACACCC	CCCATCAACTGAAAG	TTCAAGCTCATC	1860
Qy	1861	ATCAGATAGTTTCAACTT	CACAGCTTAGCTCAAC	CACTCCAGACACAAT	TATATAGTAGGCAC	1920
Db	1861	ATCAGATAGTTTCAACTT	CACAGCTTAGCTCAAC	CACTCCAGACACAAT	TATATAGTAGGCAC	1920
Qy	1921	TACCAATCCTAAACAAT	TATACGCAACAATCAAA	TACACCCCTCATCAACA	AAAATCAGAA	1980
Db	1921	TACCAATCCTAAACAAT	TATACGCAACAATCAAA	TACACCCCTCATCAACA	AAAATCAGAA	1980
Qy	1981	TCCTCAACAGCACAACCA	1999			
Db	1981	TCCTCAACAGCACAACCA	1999			

```

RESULT      2
US-09-765-271-1
; Sequence 1, Application US/09765271
; GENERAL INFORMATION:
;
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;

```

```

COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,271
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/536,784
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/961,083
FILING DATE: OCT-30-1997
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
LENGTH: 1999 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```

	Query Match	100.0%;	Score 1999;	DB 30;	Length 1999;	
	Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches	0;	Gaps
	Matches 1999;	Conservative	0;	Indels	0;	0;
Qy	1	TAAATCTACCAACAATAAATAACATTCATTGGTGACTTGGGTCTGAAGCCCGGTCAA	60			
Db	1	TAAATCTACCAACAATAAATAACATTCATTGGTGACTTGGGTCTGAAGCCCGGTCAA	60			
Qy	61	TGCCCAAAGCTAATGATATATCCCACAGATTTGGTTTAAGGCAATCGTTTTCTATCGAAGACCA	120			
Db	61	TGCCCAAAGCTAATGATATATCCCACAGATTTGGTTTAAGGCAATCGTTTTCTATCGAAGACCA	120			
Qy	121	TCGGTCTCTGCACCACAGGGGGGATGATACATCCGTATCCTGGAGGCTTCTTTGCCCA	180			
Db	121	TCGGTCTCTGCACCACAGGGGGGATGATACATCCGTATCCTGGAGGCTTCTTTGCCCA	180			
Qy	181	TCTGCAAGCAATTTCCCTCCAAGGTGGATCAACTCTCACCCAACAGTTGATTAAGTTGAC	240			
Db	181	TCTGCAAGCAATTTCCCTCCAAGGTGGATCAACTCTCACCCAACAGTTGATTAAGTTGAC	240			

ZIP: 20830
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

Db 1 TAAATCTACGCAATAAATACTCACTATGCTGACTTGGGTCTGTAACGCGCGTCAA 60
Qy 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Qy 121 TCGCTTCTCGACACAGGGGATGATACCAATCGATATCCCTGGGAGCTTCTTGGCGAA 180
Db 121 TCGCTTCTCGACACAGGGGATGATACCAATCGATATCCCTGGGAGCTTCTTGGCGAA 180
Qy 181 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCACACAGTTGATTAAGTTGAC 240
Db 181 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCACACAGTTGATTAAGTTGAC 240
Qy 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 300
Db 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 300
Qy 301 AGCGATTCAAGTAAACAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360
Db 301 AGCGATTCAAGTAAACAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360
Qy 361 GGTCTACATGCTAATGGGAATATGGGAATGGAATGCAAGCAAGCAAGCAAGCAAGCAAG 420
Db 361 GGTCTACATGCTAATGGGAATATGGGAATGGAATGCAAGCAAGCAAGCAAGCAAGCAAG 420
Qy 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGAAATGCCTCAGGCACC 480
Db 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGAAATGCCTCAGGCACC 480
Qy 481 AAACCAATATGACCCCTATTACATCCAGAACAGCCCAAGCAAGCAAGCAAGCAAGCAAG 540
Db 481 AAACCAATATGACCCCTATTACATCCAGAACAGCCCAAGCAAGCAAGCAAGCAAGCAAG 540
Qy 541 ATCTGAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAAGTCAATAC 600
Db 541 ATCTGAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAAGTCAATAC 600
Qy 601 ACCAATTACTGATGAGTACAAAGTCTCAATCAGCAAGTAAATACCCCTGCTTACATGGA 660
Db 601 ACCAATTACTGATGAGTACAAAGTCTCAATCAGCAAGTAAATACCCCTGCTTACATGGA 660
Qy 661 TAATTAACCTCAAGGAGTCAATCAAGTTGAAGAAAGCAAGGCTATAACCTACTAC 720
Db 661 TAATTAACCTCAAGGAGTCAATCAAGTTGAAGAAAGCAAGGCTATAACCTACTAC 720
Qy 721 AACTGGATGATGCTACACAAATGATAGCAAGCAAGCTCAAAACATCTGTGGGATAT 780
Db 721 AACTGGATGATGCTACACAAATGATAGCAAGCAAGCTCAAAACATCTGTGGGATAT 780
Qy 781 TTACAATACAGCAATACGTTTGGCTATCCAGACGATGAATTTGCAAGTCCGCTTACCAT 840
Db 781 TTACAATACAGCAATACGTTTGGCTATCCAGACGATGAATTTGCAAGTCCGCTTACCAT 840
Qy 841 TGTGATGTTTCAACGTTAAAGTCAATTTGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
Db 841 TGTGATGTTTCAACGTTAAAGTCAATTTGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
Qy 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACACCGGAGTGGGATCAACTATGAA 960
Db 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACACCGGAGTGGGATCAACTATGAA 960
Qy 961 ACCGATACAGACTATGCTTCCCTGCTGGAGTACGCTGCTACGATTCACCTGCTACTAT 1020
Db 961 ACCGATACAGACTATGCTTCCCTGCTGGAGTACGCTGCTACGATTCACCTGCTACTAT 1020
Qy 1021 CGTTCAGATGACCCCTATACTACCTGGGACAAATACTCTGTTTATACTGGGATAG 1080
Db 1021 CGTTCAGATGACCCCTATACTACCTGGGACAAATACTCTGTTTATACTGGGATAG 1080
Qy 1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAAATCGGAAAGCTCCAGC 1140
Db 1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAAATCGGAAAGCTCCAGC 1140

RESULT 4

US-60-029-960-81/c

; Sequence 81, Application US/60029960

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 1649

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

Qy 1141 CGTGGAACTCTAAACAAGTGGGACTCAACGCGCCAGAGACTTTCTTAAATGGTCTAGG 1200
Db 1141 CGTGGAACTCTAAACAAGTGGGACTCAACGCGCCAGAGACTTTCTTAAATGGTCTAGG 1200
Qy 1201 AATCGACTACCAAGTATTTCACTACTCAATGCAATTTCAAGTAAACACCAAGTAAATCAGA 1260
Db 1201 AATCGACTACCAAGTATTTCACTACTCAATGCAATTTCAAGTAAACACCAAGTAAATCAGA 1260
Qy 1261 CAAAAAATATGAGCAAGTAGTGAAGAATGGCTGCTTACGCTGCTTACGCTGCTTTGCAAAATGG 1320
Db 1261 CAAAAAATATGAGCAAGTAGTGAAGAATGGCTGCTTACGCTGCTTACGCTGCTTTGCAAAATGG 1320
Qy 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTTAGTATGGGAGTGAATAA 1380
Db 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTTAGTATGGGAGTGAATAA 1380
Qy 1381 AGAGTTCTCTAATGTCGGAACCTGTCGCATGAAGGAAACGACAGCCTATATGATGACCGA 1440
Db 1381 AGAGTTCTCTAATGTCGGAACCTGTCGCATGAAGGAAACGACAGCCTATATGATGACCGA 1440
Qy 1441 CATGATGAAACAGTCTTGACTTATGGAATGGAACTGGACGAAATGCCTATCTTGGCTGCC 1500
Db 1441 CATGATGAAACAGTCTTGACTTATGGAATGGAACTGGACGAAATGCCTATCTTGGCTGCC 1500
Qy 1501 TCAGGCTGTAAACAGGAACTCTTAACATATACAGACGAGGAAATTTGAAACCCACATCAA 1560
Db 1501 TCAGGCTGTAAACAGGAACTCTTAACATATACAGACGAGGAAATTTGAAACCCACATCAA 1560
Qy 1561 GACCTCTCAATTTGTAGCACCTGATGAATTTTGTGCTGCTATACGCGTAAATATTTCAAT 1620
Db 1561 GACCTCTCAATTTGTAGCACCTGATGAATTTTGTGCTGCTATACGCGTAAATATTTCAAT 1620
Qy 1621 GCGTGTATGGACAGGCTATTCTAACCGTCTGACACACCTTTAGGCAATGGCTTACGGT 1680
Db 1621 GCGTGTATGGACAGGCTATTCTAACCGTCTGACACACCTTTAGGCAATGGCTTACGGT 1680
Qy 1681 CGTGTCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGGAGCAATCCAGAA 1740
Db 1681 CGTGTCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGGAGCAATCCAGAA 1740
Qy 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTATTTTAAATATGGTCTCG 1800
Db 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTATTTTAAATATGGTCTCG 1800
Qy 1801 TTCTAGCTGAACTACCTGCTCCCAACACACCCCATCACTGAAAGTTCAAGCTCATC 1860
Db 1801 TTCTAGCTGAACTACCTGCTCCCAACACACCCCATCACTGAAAGTTCAAGCTCATC 1860
Qy 1861 ATCAGATAGTTCAACTTCACTAGCTTCACTCAACCTCCCAAGCACAATAATAGTACGAC 1920
Db 1861 ATCAGATAGTTCAACTTCACTAGCTTCACTCAACCTCCCAAGCACAATAATAGTACGAC 1920
Qy 1921 TACCAATCTCAATAATACGCAACAATAACAACCCCTGATCAACAAATACAGAAATCAGAA 1980
Db 1921 TACCAATCTCAATAATACGCAACAATAACAACCCCTGATCAACAAATACAGAAATCAGAA 1980
Qy 1981 TCTTCAACACGACACACCA 1999
Db 1981 TCTTCAACACGACACACCA 1999

; PRIOR APPLICATION NUMBER: US 09/107,433									
; PRIOR FILING DATE: 1998-06-30									
; PRIOR APPLICATION NUMBER: US 60/085,131									
; PRIOR FILING DATE: 1998-05-12									
; PRIOR APPLICATION NUMBER: US 60/051,553									
; PRIOR FILING DATE: 1997-07-02									
; NUMBER OF SEQ ID NOS: 5322									
; SEQ ID NO 1312									
; LENGTH: 2160									
; TYPE: DNA									
; ORGANISM: Streptococcus pneumoniae									
US-09-593-110-1312									
Query Match 99.4%; Score 1987.8; DB 22; Length 2160;									
Best Local Similarity 99.6%; Pred. No. 0;									
Matches 1992; Conservative 0; Mismatches 7; Indels 0; Gaps									
Qy	1	TAAAAATCTACGACAAATAAAATCAACTCATTTGCTGTGACTTGGGTCTCTGACGGCGCGCTCAA	600						
Db	159	Taaatctacgacataaaatacaactcatgtgacttgggtctctgaaacgcgcgtcaa	218						
Qy	61	TGCCCAAGCTAATCATATTCACAGATTTGGTTAAGGCAATCGTTCTCATCGAAGACCA	120						
Db	219	tgcccaagctaatgatattccacagatttgggtaaaggcaatcgttctatcgaagacca	278						
Qy	121	TCGGTCTTCTGCACACAGGGGGATTGATACATCCGCTATCTCGGAGCTTTCTTGGCGCAA	180						
Db	279	tcgcttctcgaccacaggggattgataccatccgtatcctgggagcttcttgcgcaa	338						
Qy	181	TCTGAAAGCAATTTCCCTCCAAAGTGGATCACTCTCACCCACAGCTTGATTAAGTTGAC	240						
Db	339	ctcgaaagcaattccctccaaagggtgatacactctcaactcaacagttgatlaagttgac	398						
Qy	241	TTACTTTTCAACTTCGACTTCCGACCAGACTATTCTCGTAAGGCTCAGGAAGCTTGGTT	300						
Db	399	ttactttcaacttcgacttccgaccagactattctctgtaaaggctcaggaaagcttgggtt	458						
Qy	301	AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	360						
Db	459	agcgatttcagttagaacaaaagcaaaccaagcaagaatacttgacctactataataa	518						
Qy	361	GGTCTACATGTCATTATGGGAACATATGGAATGCAGACAGCAAGCTCAAAATCTACTATG	420						
Db	519	ggctctacatgctctaattggaaactatggaaatgcagacagcagctcacaactactatgtaa	578						
Qy	421	AGACCTCAATAAATTAAGTTTACCTTCAGTTAGCCTTGCTGGCTGGAATGCCCTCAGG	480						
Db	579	agacctcaataattaaagtttaoctcagttagccttgcgtgcggaatgcctcaggcacc	638						
Qy	481	AAACCAATATGACCCCTATTTCACATCCAGAAGCAGCCCAAGACCGCCGAAACTTGGT	540						
Db	639	aaaccaatatgacccctattcaatccagaagcagcccaagacccgcgaaacttggctctt	698						
Qy	541	ATCTGAAATGAAAATCAAGGCTACATCTCTGCTGAACAGTAGTAGAAAGCAGTCAATAC	600						
Db	699	atctgaaatgaaaaatcaaggctacatctctgtgaacagtagaagagcagtcacat	758						
Qy	601	ACCAATTTACTGATGGACTACAAAGTCTCAATCAGCAAGTAATACCTTGCCTTACATGA	660						
Db	759	accaattactgatgggtacaaagctcacaatcagaagtaattacccctgccttacatgga	818						
Qy	661	TAATTACTCAAGCACTCATCAATCAAGTTGAAGAAACAGGCTATTAACCTACTCAC	720						
Db	819	taa ttaoctcaagggaagtcataatcaag tgaagaagaacaagccta taaactactcac	878						
Qy	721	AACTGGGATGGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAAACATCTCTGGG	780						
Db	879	aactgggatggatgctctacacaaatgtagacaaagagctcacaacacatctgtgggata	938						
Qy	781	TTACAATACAGAGCAATACGTTGCCTATCCAGACGATGAATTCGAAGTCGCTTACCAT	840						
Db	939	ttacaa tacagacaa lacgttaccctaccagacagatgaattcaagtcgcttccact	998						

```

RESULT 6
US-09-583-110-1312
; Sequence 1312, Application US/09583110
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26

```


QY 841 TGTGTGTTCTTAACGGTAAAGTCAATGCCCCAGCTAGGAGCAGCGCATCAGTCAAGTAA 900
Db 999 tgttgatgtttctcaacggttaaagtcattgcccagctagagcagcgcacatcagtaagtaa 1058
QY 901 TGTTCCTTCGGAAATTAACCAAGCAGTACAGAAACACCGGACTGGGGATCAACTATGAA 960
Db 1059 tgtttcttcggaattcaaccaagcagtagaacaacacccgactggggatcaactatgaa 1118
QY 961 ACCGATCACAGACTATGCTCCTGCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT 1020
Db 1119 accgatacacagactatgctcctgcttggtagtcaggtatctacgattcaactgctactat 1178
QY 1021 CGTTCAGATGAGCCCTATAACTACCCCTGGGACAAATACCTCCGTTTATAACTGGGATAG 1080
Db 1179 cgttcacgatgagccctataactacccctgggacaaataccctgtttataactgggatag 1238
QY 1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACATCGGAAACGTCCCGC 1140
Db 1239 gggctactttggcaacatcaccttgcaatcacgacctgcaacatcgcaaacgctccagc 1298
QY 1141 CGTGAACCTCTAAACAAGTCGAGCTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG 1200
Db 1299 cgtgaaactctaaacaaagtcgactcaacccgccaagactttcctaataatggtctcgg 1358
QY 1201 AATCGACTACCCAAAGTATTCACTACTCAATGCCATTTTCAAGTAAACACCAACCAATCAGA 1260
Db 1359 aatcgactacccaagtattcaactactcaaatgccatttcaagttaacacacacgaatcaga 1418
QY 1261 CAAAAATATGGAGCAAGTAGTCAAAAGATGGCTGTCTAGCTGCCTTTGCCAATGG 1320
Db 1419 caaaaaatgtagcaagtagtgaagaatggctgtctgctacgctgccttggcaaatgg 1478
QY 1321 TGAACCTACTATAAACCAACATGTATCCATAAAGTCGTCTTTAGTATGGGAGTGAATA 1380
Db 1479 tgaacctactataaaccaatgtatccataaagtcgtctttagtgatggagtgaaaa 1538
QY 1381 AGAGTTCTTAATGTCGGAACGTGTCGCATGAAGGAACAGACGCCTATATGATGACCGA 1440
Db 1539 agagttctctaattgcggaactgtgccaatgaaggaaacgacgctatatgatgaccga 1598
QY 1441 CATGATGAACACAGCTCTGACTTATGSAATGACCAATGCTACTCTGCTGGCTCCC 1500
Db 1599 caugaagaacacgctctgacttatggaactggaactggaacgaatgctatctgctggctccc 1658
QY 1501 TCAGGCTGTAAACAGGAACCTCTAACTATACAGACGAGGAAATGAAACACACATCAA 1560
Db 1659 tcaggctgttaaacaggaacctctaactatacagacgaggaaaattgaaacacacataaa 1718
QY 1561 GACCTCTCAATTTGTAGCACCTGATGAATATTTGCTGGCTATACGCGTAAATATTTCAAT 1620
Db 1719 gacctctcaattgtagcacctgatgaactatttggctatacgcgcgtataatattcaat 1778
QY 1621 GCGTGATGACAGGCTATTCTAACCGTCTGACACACTGTAGGCAATGGCCTTACGGT 1680
Db 1779 ggcgtatggacaggctaattcaacgctgacacacactgtaggcaatggcctacggt 1838
QY 1681 CGGTGCAAGTTTACCGCTCTATGATGACCTACCTGCTGAGGAAGCAATCCAGAGA 1740
Db 1839 cgctgcaaaagtttacccgctctatgatgacctatctgtctgaaggaaagcaatccagagga 1898
QY 1741 TTGGAATATACGAGGGGCTCTACAGAAATGGAGAAATTCGATTTTAAATGGTGCTCG 1800
Db 1899 ttggaataaccagagggctctacagaaatggagaattcgatatttaaaaaatggtgctcg 1958
QY 1801 TTCTAGTGGAACTACCTGCTCCACAAACCCCTCACTCAACTGAAGTTCAGCTCATC 1860
Db 1959 ttctacgtggaactcaactcgcctccacaacaccccccatcaactgaaagttcaagctcatc 2018
QY 1861 ATCAGATAGTTCAACTTACAGTCTAGCTCAACACTCCCAAGCACAAATAATAGTACGAC 1920
Db 2019 atcagatagttcaacttcaagcttagctcaaccactcccaagcacaaaaataatagtagcagc 2078

QY 1921 TACCAATCCTTAACAATAATACGCAACAATAACAACCCCTGATCAACAATAATCAGAA 1980
Db 2079 taccatctcaacaataataacgcaacaatacaacccctgatcaacaataatcagaa 2138
QY 1981 TCCTCAACCAAGCAACCA 1999
Db 2139 tctcaaccagcacaacca 2157
RESULT 7
US-09-107-433-1102
; Sequence 1102, Application US/09107433
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIA
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2166 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1..2166
; SEQUENCE DESCRIPTION: SEQ ID NO: 1102:
US-09-107-433-1102
Query Match 99.4%; Score 1987.8; DB 15; Length 2166;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 TAAATCTACGACAATAAAAAATCAACTCATTTGCTGACTTGGGTCTGAAACGCGCGTCAA 60
Db 165 TAAATCTACGACAATAAAAAATCAACTCATTTGCTGACTTGGGTCTGAAACGCGCGTCAA 224
QY 61 TGCCCAAGCTAATGATATTCCCAACAGATTGGTTAGGCAATCGTTTCTATCGAAGACCA 120
|||||

Db 225 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 284
 QY 121 TCCTCTTCGACACACAGGGGATGATACCATCCGTAATCTCGGAGCTTCTTCGCGAA 180
 Db 285 TCCTCTTCGACACACAGGGGATGATACCATCCGTAATCTCGGAGCTTCTTCGCGAA 344
 QY 181 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCCCAACAGTTGATTAAAGTTGAC 240
 Db 345 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCTCAACAGTTGATTAAAGTTGAC 404
 QY 241 TTTACTTTTCACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGAGAGCTTGGTT 300
 Db 405 TTTACTTTTCACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGAGAGCTTGGTT 464
 QY 301 AGCGATTCTAGTATAGAACAAAACCAAGCAAGCAAGAAATCTTGACCTACTATATAATAAA 360
 Db 465 AGCGATTCTAGTATAGAACAAAACCAAGCAAGCAAGAAATCTTGACCTACTATATAATAAA 524
 QY 361 GGTCTACATGTCTAATGGGAACCTATGGAATGCAGACAGAGCTCAAAACCTACTATGATAA 420
 Db 525 GGTCTACATGTCTAATGGGAACCTATGGAATGCAGACAGAGCTCAAAACCTACTATGATAA 584
 QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGTGAATGGCTCAGGACAC 480
 Db 585 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGTGAATGGCTCAGGACAC 644
 QY 481 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGACCGCGGAAACTTGGCTCT 540
 Db 645 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGACCGCGGAAACTTGGCTCT 704
 QY 541 ATCTGAATGAAATCAAGGCTTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
 Db 705 ATCTGAATGAAATCAAGGCTTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 764
 QY 601 ACCAATTTACTGATGAGTACAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGGA 660
 Db 765 ACCAATTTACTGATGAGTACAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGGA 824
 QY 661 TAATTAACCTCAGAGAGTATCAATCAAGTTGAAGAGAAACAGGCTTAACCTACTAC 720
 Db 825 TAATTAACCTCAGAGAGTATCAATCAAGTTGAAGAGAAACAGGCTTAACCTACTAC 884
 QY 721 AACTGGGATGGATGCTTACACAAATAGACCAAGAGCTCAAAACATCTGTGGGATAT 780
 Db 885 AACTGGGATGGATGCTTACACAAATAGACCAAGAGCTCAAAACATCTGTGGGATAT 944
 QY 781 TTACAATACAGAGATAGTTGCTTATCCAGACGATGAATTCGAAGTGGCTTCTACCAT 840
 Db 945 TTACAATACAGAGATAGTTGCTTATCCAGACGATGAATTCGAAGTGGCTTCTACCAT 1004
 QY 841 TGTGTATGTTTCTTAACGGTAAAGTCAATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
 Db 1005 TGTGTATGTTTCTTAACGGTAAAGTCAATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 1064
 QY 901 TGTTCCTCTCGGAATTAACCAAGCAGTAGAACAACACCGAGCTGGGGATCAACTATGAA 960
 Db 1065 TGTTCCTCTCGGAATTAACCAAGCAGTAGAACAACACCGAGCTGGGGATCAACTATGAA 1124
 QY 961 ACCGATACAGAGATAGTCTGCTTGGAGTACGGTGTCTAGGATTCAGTTCAGTCTACTAT 1020
 Db 1125 ACCGATACAGAGATAGTCTGCTTGGAGTACGGTGTCTAGGATTCAGTTCAGTCTACTAT 1184
 QY 1021 CGTTACAGTACGACCTTAACCTACCTCGGACAAATACCTCTGTTTATAAAGTGGGATAG 1080
 Db 1185 CGTTACAGTACGACCTTAACCTACCTCGGACAAATACCTCTGTTTATAAAGTGGGATAG 1244
 QY 1081 GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACAAATCGCGAAACGTCGCCAG 1140
 Db 1245 GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACAAATCGCGAAACGTCGCCAG 1304
 QY 1141 CGTGGAACTCTAAACAAGGTCGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG 1200
 Db 1305 CGTGGAACTCTAAACAAGGTCGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG 1364

QY 1201 AATCGACTACCCAAAGTATTCACCTACTCAATGCCATTTCAAGTAACACACCAACCAATCAGA 1360
 Db 1365 AATCGACTACCCAAAGTATTCACCTACTCAATGCCATTTCAAGTAACACACCAACCAATCAGA 1424
 QY 1261 CAAAAAATATGGAGCAAGTAGTGAAGAGTGGCTGCTGTAGCGTTCGCTTTGCAAAATGG 1320
 Db 1425 CAAAAAATATGGAGCAAGTAGTGAAGAGTGGCTGCTGTAGCGTTCGCTTTGCAAAATGG 1484
 QY 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCTGCTTTAGTGTAGTGGAGTGAATA 1380
 Db 1485 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCTGCTTTAGTGTAGTGGAGTGAATA 1544
 QY 1381 AGAGTTCTCTAATGTGCGGAACCTGCTGCCATGAAGGAAACGACAGCCTATATGATGACCGA 1440
 Db 1545 AGAGTTCTCTAATGTGCGGAACCTGCTGCCATGAAGGAAACGACAGCCTATATGATGACCGA 1604
 QY 1441 CATGATGAAACAGTCTTGACTTTATGGAATGGAACTGGACGAAATGCTTATCTTGGCTGCC 1500
 Db 1605 CATGATGAAACAGTCTTGACTTTATGGAATGGAACTGGACGAAATGCTTATCTTGGCTGCC 1664
 QY 1501 TCAGGCTGTAAACAGGAAACCTCTAAGTATATACAGACGAGGAAATTTGAAACCAACATCAA 1560
 Db 1665 TCAGGCTGTAAACAGGAAACCTCTAAGTATATACAGACGAGGAAATTTGAAACCAACATCAA 1724
 QY 1561 GACCTCTCAATTTGTAGCACCTGATGAACCTATTTGCTGGCTATACGCGTAAATATTTCAAT 1620
 Db 1725 GACCTCTCAATTTGTAGCACCTGATGAACCTATTTGCTGGCTATACGCGTAAATATTTCAAT 1784
 QY 1621 GGCTGTATGACAGGCTTATTAACCGTCTGACACACCTGTAGGCAATGGCCTTACGGT 1680
 Db 1785 GGCTGTATGACAGGCTTATTAACCGTCTGACACACCTGTAGGCAATGGCCTTACGGT 1844
 QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTTACCTGCTGAAGGAGCAATCCAGAA 1740
 Db 1845 CGCTGCCAAAGTTTACCGCTCTATGATGACCTTACCTGCTGAAGGAGCAATCCAGAA 1904
 QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 1800
 Db 1905 TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 1964
 QY 1801 TTCTAGTGGAACTACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
 Db 1965 TTCTAGTGGAACTACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 2024
 QY 1861 ATCAGATAGTTCAACTTCACAGTCTACAGTCTCAACCTCCAGCAGCAAAATTAATAGTAGC 1920
 Db 2025 ATCAGATAGTTCAACTTCACAGTCTACAGTCTCAACCTCCAGCAGCAAAATTAATAGTAGC 2084
 QY 1921 TACCAATCTCAACAATAATACGCAACAATCAATAACAACCCCTGATCAACAAATCAGAA 1980
 Db 2085 TACCAATCTCAACAATAATACGCAACAATCAATAACAACCCCTGATCAACAAATCAGAA 2144
 QY 1981 TCCTCAACAGCAGCAACCA 1999
 Db 2145 TCCTCAACAGCAGCAACCA 2163

RESULT 8

PCT-US02-03987-9325
 ; Sequence 9325, Application PC/TUS0203987
 ; GENERAL INFORMATION:
 ; APPLICANT: Elitra Pharmaceuticals, Inc.
 ; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit
 ; TITLE OF INVENTION: Proliferation
 ; FILE REFERENCE: ELITRA.028VPC
 ; CURRENT APPLICATION NUMBER: PCT/US02/03987
 ; CURRENT FILING DATE: 2002-02-02
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; NUMBER OF SEQ ID NOS: 15811
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9325

[illegible]


```
QY 1441 CATGATGAAACAGTCTTGACTTATGAACTGGACGAAATGCCATCTTGCTGGCTCCC 1500
|||||
Db 1599 catgatgaaacagctcttgagttatgaaactgacgaatgctatctgttgctgctccc 1658
QY 1501 TCAGGCTGTAAACAGGAACTTAACTATACAGACGAGGAAATGAAACACATCAA 1560
|||||
Db 1659 tcaggctgtgtaaacaggaaacctctaactatacagacgaggaatgaaacacatcaa 1718
QY 1561 GACCTCTCAATTTGTAGCACCTGATCAACTATTTGCTGCTATACGCGTAAATATCAAT 1620
|||||
Db 1719 gacctctcaattttagcaccctgagaaactattctgctgctacacgtaaatcaat 1778
QY 1621 GCGTGTATGACAGGCTTATTAACCGTCTGACACCACTTGTAAGCAATGGCCTTACGGT 1680
|||||
Db 1779 ggcgtgatggacaggtattcttaaccgtctgcacacacttgtagcaatggcctacggt 1838
QY 1681 CGCTGCAAAAGTTACCGCTCTATGATGACCTTACCTGCTGTAAGGAAGCAATCCAGA 1740
|||||
Db 1839 cgctgcgaagtttacgcctctatgatgacctacctgctggaaggagcaatccagagga 1898
QY 1741 TTGGAATATACGAGAGGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAATGGTCTCG 1800
|||||
Db 1899 ttggaatataccagagggctctacagaaatggagaatcgtaatttaaaaatggtgctcg 1958
QY 1801 TTCTAGCTGGAACTACCTGCTGCCACAACAACCCCATCAACTGTAAGTTCAAGCTCATC 1860
|||||
Db 1959 ttctacgtggagtcacctgctccacaacaaccccccatcaactgaaagtccaagctcatc 2018
QY 1861 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCAACGACACAATAATAGTAGC 1920
Db 2019 atcagatagttcaacttcacagttcagctcagctcaaccactccaagcacaaataatagtagc 2078
QY 1921 TACCAATCTCAACAATAATACGCAACAATAAATCAATCAACCCCTGATCAACAATAATCAGAA 1980
Db 2079 taccaatctcaacaataatcagcaacaatacaatacaacccccctgatacaacaataatcagaa 2138
QY 1981 TCCTCAACCAAGCACACCA 1999
|||||
Db 2139 tctcaaccagcacacaacca 2157
```

```
RESULT 10
US-10-072-851-9325
; Sequence 9325, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Twack, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits G
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9325
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2160)
US-10-072-851-9325

Query Match 99.1%; Score 1981.4; DB 37; Length 2160;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTAGCAATAAATAAATCAACTCATCTGCTGACTTGGGTTCTGAACCGCGCTCAA 60
Db 159 taaatctcagcaataaataaatacaactcatctgctgacttgggttctgaacgcgcgtcaa 218
QY 61 TGCCCAAGCTAATGATATTCCCAACAGATTGGTTAAGGCAATCGTTTCTATCAAGACCA 120
Db 219 tgcccaagctaatgatatattcccaacagatttgggttaaggcaatcgtttctatcagaacca 278
QY 121 TCGCTTCTCGACCAAGGGGATTGATACCATCCGTATCTCTGGAGCTTCTTTCGCGCAA 180
Db 279 tcgcttcttcgaccacaggggattgataccatccgtatcctctggagcttcttcgcgcaa 338
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGGATCACTCTCACCCAAACAGTTGATTAAGTTGAC 240
Db 339 tctgcaagcaattcctcccaagggtggtacagctcctcaactcaacagttgattaagttgac 398
QY 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 399 ttacttttcaacttcgacttcgaccagactattctcgtaaagctcaggaagcttgggt 458
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
Db 459 agcgatttcagttagaaacaaaaagcaaccaagcaagaatcttgacctactatataataa 518
QY 361 GGTCTACATCTCTAATGGGNACTATGGAATGCAGACAGAGCTCAAAACTACTATGGTAA 420
Db 519 ggtctacatgtctaattgggaactatggaatgcagacagcagctcaaaactactatgtaa 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTTCGGCTGGGAATGCTCAGACACC 480
Db 579 agacctcaataatttaagtttacctcagtttagcttgccttgccttgcctcaggcacc 638
QY 481 AAACCAATATGACCCCTATTACATCCAGAACGAGCCCAAGAGCCGCCGAACTTGGTCTT 540
Db 639 aaaccaatatgacccctattcacatccagaagcagcccaagacccgcgaacttgcctt 698
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGACAGATGAGAAACGATCAATAC 600
Db 699 atctgaaatgaaaaatcaagggtcacatctctgctgacagatgagaaagcagtcataac 758
QY 601 ACCAATTACTGTAGGACTACAAAGTCTCAAACTCAGCAAGTAATTAACCTGCTTACATGA 660
Db 759 accaattactgtatgggtcacaagttcctcaaatcagcaagtaattacccctgctcacaat 818
QY 661 TAATTACCTCAGGAAGTTCATCAATCAAGTTGAAGAAAGAAACAGGCTATAACCTACTCAC 720
Db 819 taattacctcaggaagtcatcaatcaagttgaagaagaacaggtcataaactactcac 878
QY 721 AACTGGGATGGATGCTACACAAATGTAGACCAAGAGCTCAAAACATCTGCGGATAT 780
Db 879 aactgggatggatgctacacaaatgtagaccaagagctcaaaaacatctgtgggatat 938
QY 781 TTACAATACAGACGAATACGTTGCTTATCCAGACGATGAATTCGAAGTCTGCTTCTACCAT 840
Db 939 ttacaatacagacgaataacgttgcctatccagacagatgaattgcagtgctcctaccat 998
QY 841 TGTGTATGTTTCTTAACGGTAAAGTCAATTCGCCAGCTAGGAGCAGCGCATCAGTCAAGTAA 900
Db 999 tgtgtatgtttctaaacggtaaaagtcaattgtcccaagctaggagcagcgcacatcaagtaa 1058
QY 901 TGTTTCTTGGGAATTAACCAAGCAGTAGAAACAAACCCGACTGGGGATCAACTATGAA 960
Db 1059 tgtttcttcgggaattaaaccaagcagtagaacaacccgcgactggggatcaactatgaa 1118
```



```

QY 961 ACCGATCACAGACTATGCTCCTGCTGGAGTACGGTGTCTAGATTCAACTGCTACTAT 1020
Db 1119 accgatcacagactatgctcctgctggagtagcgtgtctacgagtaactgccaactat 1178
QY 1021 CGTTCAGATGAGCCCTATAACTACCTCGGCAATAACTCTCTGTTTATAACTGGGATAG 1080
Db 1179 cgttcagatgagccctataactacccctggacaataccctgttataaactgggatag 1238
QY 1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACATCGGAACGTCCACAG 1140
Db 1239 gggctactttggcaaacatcaccttgcaatcgccctgcaacatcggaacgtccacag 1298
QY 1141 CBTGGAACCTCTAAACAAGCTCGAGTCAACCGCGCAAGACTTCTCTAAATGGTCTAG 1200
Db 1299 cbtggaacctctaaacaagctcgagctaacccgcgaagacttctctaaatggctag 1358
QY 1201 ATTCGACTACCCAAAGTATTCACTACTACCAATGCAATGCAATGCAACACCAACGAAATCAGA 1260
Db 1359 aatcgactacccaagtattcaactactcaaatgccatttcaagtaacacacacgaatcaga 1418
QY 1261 CAAAAAATATGGAGCAAGTAGTCAAAAGATGGCTGCTGTAGCTGCTTTCCTCAATGG 1320
Db 1419 caaaaatatggagcaagtagtcaaaagatggctgctgttagcgtgcttgcgaatgg 1478
QY 1321 TGGAACTTACTATAAACCAATGATATCCATAAAGTCGTCTTTAGTGTGGGAGTGAAAA 1380
Db 1479 tggaaacttactataaaccaatgatataccataaagtcgtctttagtgtgggagtgaaaa 1538
QY 1381 AGAGTCTCTAATGTGCGAAGTCGTGCCATGAAGGAAACGACAGCCATATATGATGACCGA 1440
Db 1539 agagttctctaaatgtcggaactcgtgccatgaagaaacgacagccatataatgatgaccca 1598
QY 1441 CATGATGAAACAGTCTTGACTTATGAACTGGACCAATGCTATCTTGTGGCTCCC 1500
Db 1599 catgagaaacagtccttgacttatgaaactggacccaatgctatcttgtgtggctccc 1658
QY 1501 TCAGGCTGTGTAACAGGAACTCTAACTATACAGACGAGGAAATGAAACACCATCAA 1560
Db 1659 tcaggctgtgtaaacaggaaaccttaactatacagacgaggaattgaaacacacatcaa 1718
QY 1561 GACCTCTCAATTTGTAGCACCCTGATGAATTTTGTGCTATACCGTAAATATCAAT 1620
Db 1719 gacctctcaattttagcaccttgatgaactatttggctgtatacgcgtaaatattcaat 1778
QY 1621 GGCTGTATGACAGGCTATTCTAACCGTCTGACACCACTTGAGCAATGGCTTACGGT 1680
Db 1779 ggcctgtatggacaggctattctaaccgctctgacaccacttgaggcaatggccttacggt 1838
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTGTAAGGAAGCAATCCAGAAGA 1740
Db 1839 cgcgtgcaaaagtttacccgctctatgatgacctacctgtctgaaggaagcaatccagagga 1898
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGTATTTAAATAATGGTCTCG 1800
Db 1899 ttggaataacagagagggtctacagaaatggagaattcgtattttaaataatgggtctcg 1958
QY 1801 TTCTACGTGGAATCACTACCTGCTCCACAACACCCCATCAACTGAAAGTTCAAGCTCATC 1860
Db 1959 ttctacgtggagctaccctgctccacaacaccccatcaactgaaagtccaagctcatac 2018
QY 1861 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCCCAAGCACAAATATAGTACGAC 1920
Db 2019 atcagatagtccaacttcacagcttagctcaaccactccaagcacaaataatagtcagac 2078
QY 1921 TACCATCTCTACAAATATACGCAACAATCAAAATACACCCCTGATCAACAAATCAGAA 1980
Db 2079 taccatctctacaataatcacgacaatacaaatatacaaccctgatcaacaataatcagaa 2138
QY 1981 TCTCTAACCCAGCACAAACCA 1999
Db 2139 tctctaacccagcacaaacca 2157

```

```

RESULT 11
US-60-061-998-596/c
; Sequence 596, Application US/60061998
; GENERAL INFORMATION:
; APPLICANT: LAGACE, ROBERT E.
; APPLICANT: CORLEY, NEIL C.
; APPLICANT: RUSSO, FRANK D.
; APPLICANT: HENN, AMY L.
; APPLICANT: HENH, JOE D.
; APPLICANT: FINNEY, GREGORY L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE
; NUMBER OF SEQUENCES: 797
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/061,998
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0006-2P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 596:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10333 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: SPN1c601
; US-60-061-998-596

```

```

Query Match 99.1%; Score 1981.4; DB 45; Length 10333;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATAAATAAATACTATGCTGACTTGGGTTCTGAACGGCGCTCAA 60
Db 9718 TAAATCTACGACATAAATAAATACTATGCTGACTTGGGTTCTGAACGGCGCTCAA 9659
QY 61 TGCCCAAGCTAATGATATTCGCCACAGATTGGTTAAGGCAATCGTTTCTATCGAACACCA 120
Db 9658 TGCCCAAGCTAATGATATTCGCCACAGATTGGTTAAGGCAATCGTTTCTATCGAACACCA 9599
QY 121 TCGCTTCTTCGACCACAGGGGGATTGATACCATCCCTGCGAGCTTCTTCTGCGCAA 180
Db 9598 TCGCTTCTTCGACCACAGGGGGATTGATACCATCCCTGCGAGCTTCTTCTGCGCAA 9539
QY 181 TCTGCAACCAATTCCTCCAGGTGGATCACTCTACCCCAACAGTTGATTAAGTTGAC 240
Db 9538 TCTGCAACCAATTCCTCCAGGTGGATCACTCTACCCCAACAGTTGATTAAGTTGAC 9479
QY 241 TTACTTTTCAACTTCGACTTCGACCTTCGACCCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 9478 TTACTTTTCAACTTCGACTTCGACCTTCGACCCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 9419

```


ATTORNEY/AGENT INFORMATION:
 NAME: CERRONE, MICHAEL C.
 REGISTRATION NUMBER: 39,132
 REFERENCE/DOCKET NUMBER: PM-0009-2 P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 855-0555
 TELEFAX: (650) 845-416
 INFORMATION FOR SEQ ID NO: 596:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9845 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: genomic DNA
 IMMEDIATE SOURCE:
 CLONE: SPN2c598
 US-60-068-175-596

Query Match 81.6%; Score 1631.2; DB 45; Length 9845;
 Best Local Similarity 88.8%; Pred. No. 0;
 Matches 1776; Conservative 0; Mismatches 223; Indels 1; Gaps 1;

QY 1 TAAATCTACGACATATAAATCAACTCATGCTGACCTGGGTTCTGAACCGCGGTCAA 60
 DB 9734 TAAATCTACGACATATAAATCAACTCATGCTGACCTGGGTTCTGAACCGCGGTCAA 9675

QY 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA 120
 DB 9674 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA 9615

QY 121 TCGCTTCTTCGACACAGGGGATATGATACCATCGTATCTCTGGGAGCTTTCTTCGCAA 180
 DB 9614 TCGCTTCTTCGACACAGGGGATATGATACCATCGTATCTCTGGGAGCTTTCTTCGCAA 9555

QY 181 TCTGCAAGCAATTCCTCCAGGTGGATCAACTCTCACCCAAACAGTTGATTAAGTTGAC 240
 DB 9554 TCTGCAAGCAATTCCTCCAGGTGGATCAACTCTCACCCAAACAGTTGATTAAGTTGAC 9495

QY 241 TTACTTTTCAACTTCGACCTTCGGACAGACTATTTTCTCGTAGGCTCAGGAAGCTTGGTT 300
 DB 9494 TTACTTTTCAACTTCGACCTTCGGACAGACTATTTTCTCGTAGGCTCAGGAAGCTTGGTT 9435

QY 301 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA 360
 DB 9434 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA 9375

QY 361 GGTCTACATGCTTAATGGAACTATGGAATGCAGACAGAGCTCAAACTACTATGGTAA 420
 DB 9374 GGTCTACATGCTTAATGGAACTATGGAATGCAGACAGAGCTCAAACTACTATGGTAA 9315

QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGGCTGGTGAATGCCCTCAGGCACC 480
 DB 9314 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGGCTGGTGAATGCCCTCAGGCACC 9255

QY 481 AAACCAATATGACCCCTATTACATCCAGAACGCCAACCAAGACCCGCAAACTTGGTCTT 540
 DB 9254 AAACCAATATGACCCCTATTACATCCAGAACGCCAACCAAGACCCGCAAACTTGGTCTT 9195

QY 541 ATCTGAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC 600
 DB 9194 ATCTGAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC 9135

QY 601 ACCAATTAAGTACGACATCAAGTCTCAATCAGCAAGTATTAACCTT-GCTTACATGG 659
 DB 9134 ACCAATTAAGTACGACATCAAGTCTCAATCAGCAAGTATTAACCTTACCTTACATGG 9075

QY 660 ATAATTACCTCAAGGAAGTCATCAATCAAGTTGAAGAAGAAACAGGCTATTAACCTACTCA 719
 DB 9074 ATAATTACCTCAAGGAAGTCATCAATCAAGTTGAAGAAGAAACAGGCTATTAACCTACTCA 9015

QY 720 CAATCGGATGATGCTTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATA 779
 DB 779 CAATCGGATGATGCTTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATA 7875

DB 9014 CTACTGGGATGGATGTTTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATA 8955

QY 780 TTTACAATACAGACGAATACGTTGCTTATCCAGACAGATGAATTGCAAGTTCGCTTCTACCA 839

DB 8954 TCTACAACCTCGATCAATACGCTCTTACCTGACGATGATTTGCAAGTTCGCTTCTACGG 8895

QY 840 TTGTTGATGTTTCTAACGGTAAAGTCAATGCCACGCTAGGAGCAGCCCATCAGTCAAGTA 899

DB 8894 TCGTAGATGTTTCAAAATGTTAAAGTCATCGCCCAACTTGGAGCTCGTCAACCAAGCAATGA 8835

QY 900 ATGTTTCTTCGGAATTAACCAAGCAGTAGAACAACACCGGACCTGGGATCACTATGA 959

DB 8834 ACGTTTCAATTTGGTACCAACCAAGCTGGAACAACCAATCGTACGTTGCTTCTGCTATGA 8775

QY 960 AACGATCACAGACTATGCTCTGCTGTTGGAGTAGCGGTCTACGATTCACATCAACTGCTACTA 1019

DB 8774 AACCAATCACGATTTATGCACCTGCCATAGATAGCGGTGTTATGATTCCTCACTGCACTA 8715

QY 1020 TCGTTACAGATGAGCCCTTATACTACCTGGGACAAATACTCCTCTTTATTAACCTGGGATA 1079

DB 8714 TGGTTAATGATATTCCTTATAACTATCCGGGAACAAGCACACCTCTCTACAACTGGGATA 8655

QY 1080 GGGGTACTTTGGCAACATCACCCTTGCATAGCCCTCGCAACATCGGAACGTCCTCAG 1139

DB 8654 GAGCATATTCGGTAAATATTAATCTCTGCAATATGCTCTTCAACATCAGCAAAATGTCACAG 8595

QY 1140 CGGTGAAACTCTAAACAAGGTCGGACTCAACCGCGCAAGACTTTCCTAAATGCTGAG 1199

DB 8594 CGTTGAGACTTTGAATTAAGTTCGGCTTAGATAGAGCTTAAACCTTCTTAATGCTCTTG 8535

QY 1200 GAATCGACTACCCAAAGTATTCACCTACTCAATGCCATTTCAAGTAAACACCAACCAAGTACAG 1259

DB 8534 GTATCGACTATCAAGCATGCAATTTGCAACGCCATTTCAAGTAAACACCAAGTACTA 8475

QY 1260 ACAAAAATATGGACCAAGTAGTGAAGAAGTGGCTGCTGCTTACGCTGCCTTTGCAATG 1319

DB 8474 ATAAACAATCGGACCAAGTAGTGAAGAAGTGGCTGCTGCTTATGCTGCTTTGCAATG 8415

QY 1320 GTGGAACTTACTATAAACAATGATATATCCATAAAGTCTGCTTTAGTATGGAGTGA 1379

DB 8414 GTGGCACTTACTATAAACAATGATATATCCATAAAGTCTGCTTTAGTATGGAGTGA 8355

QY 1380 AAGAGTCTCTAATGTCGGAACCTCGTCCATGGAAGAAACGACGCCCTATATGATGACGG 1439

DB 8354 AAGAGTCTCTAATGTCGGAACCTCGTCCATGGAAGAAACGACGCCCTATATGATGACGG 8295

QY 1440 ACATGATGAACACAGCTCTTGACTTATGGAAGTGGAGAAATGCCTTATCTTGGCTGCC 1499

DB 8294 ACATGATGAACACAGCTCTTGACTTATGGAAGTGGGAGCTTATCTTCTTGGCTGCC 8235

QY 1500 CTCAGGCTGGTAAACAGGAACCTCTAACTATACAGACGAGGAAATTTGAAACACCATCA 1559

DB 8234 CTCAGCTGGTAAACAGGAACCTCTAACTATACAGATGAGGAAGTTGAAACACCATCA 8175

QY 1560 AGACCTCTCAATTTGTAGCAGCTGATGAATTTGCTGGCTATACGGCTAAATATTCAT 1619

DB 8174 AGAACACTGGCTTATGCTCCAGATGAAATGTTTGGTGGTATATCTCTGTAAGTATCTA 8115

QY 1620 TGGCTGATGACAGGCTTATCTAACCGCTGACACACCTTGTAGGCAATGSCCTTACGG 1679

DB 8114 TGGCTGATGACAGGCTTATCTGAATCGTTTAACTCCTATCGTTGGAGATGGTTTCTCCT 8055

QY 1680 TCGCTGCAAAAGTTTACCGCTTATGATGACCTACTGCTGAAGGAGCAATCCAGAG 1739

DB 8054 TTGACGCTTAAGTTTATCGCTCAATGATACGATATCTATCAGAAATACTCATCCAGAG 7995

QY 1740 ATTGGAATATACAGAGGGCTCTACAGAAATGGAATTCGATTTAAAAATGTTGCTC 1799

DB 7994 ACTGGCAGTGCACAGGACTTTTCAGAAACGGGAAATTTGATTTCAAAATGAGGCTC 7935

QY 1800 GTTCTACGTGGAACCTGCTGCTCCACAAACACCCCTCACTCAACTGAAAGTTCAAGCTCAT 1859

DB 7934 GCCCAATATGGACTGAACCCCTCTACTCAACAAATCTCAACAGCTGAAAGTTCAAGCTCAT 7875

APPLICANT: Stodola, Robert
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
NUMBER OF SEQUENCES: 713
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/14436
FILING DATE: 15-AUG-1997
CLASSIFICATION:
PRIOR APPLICATION NUMBER: 60/024,022
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50533
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 2172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
PCT-US97-14436-139

Query Match 40.5%; Score 808.6; DB 1; Length 2172;

Best Local Similarity 99.4%; Pred. No. 5.1e-218;

Matches 822; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1174 CCCCAAGACTTTCCTAAATGGTCTAGGAATCGACTACCCAACTATTCTACTCTCAAAATGC 1233
DB 1 CCCCAAGACTTCCCTAAATGGTCTAGGAATCGACTACCCAACTATTCTACTCTCAAAATGC 60
QY 1234 CATTTCAGTAACACACCGCAATCAGACAAAAATATGAGCAAGTAGTGAAGATGCG 1293
DB 61 CATTTCAGTAACACACCGCAATCAGACAAAAATATGAGCAAGTAGTGAAGATGCG 120
QY 1294 TGCTGCTTACGCTGCTTGGCAATGGTGGAACTTACTATTAACCAATGTATATCCATAA 1353
DB 121 TGCTGCTTACGCTGCTTGGCAATGGTGGAACTTACTATTAACCAATGTATATCCATAA 180
QY 1354 AGTCGCTCTTACTGATGGAGTGAAGAGTCTCTAATGTCGGAACCTCGGCCATGAA 1413
DB 181 AGTCGCTCTTACTGATGGAGTGAAGAGTCTCTAATGTCGGAACCTCGGCCATGAA 240
QY 1414 GGAACGACAGCCTATATGATGACCGACATGATGAAAAACAGTCTTGAAGTGAAGTGG 1473
DB 241 GGAACGACAGCCTATATGATGACCGACATGATGAAAAACAGTCTTGAAGTGAAGTGG 300
QY 1474 AGC-NAATGCCCTATCTTGGTGGCTCCCTCAGGCTGGTAAACAGGAACTCTTAATACTATA 1532
DB 301 ACGAAAAATGCCCTATCTTGGTGGCTCCCTCAGGCTGGTAAACAGGAACTCTTAATACTATA 360
QY 1533 CAGACGAGGAATTTGAAAAACCATCATCAAGACCTCTCAATTTGTAGCACTCTATGAACATAT 1592
DB 361 CAGACGAGGAATTTGAAAAACCATCATCAAGACCTCTCAATTTGTAGCACTCTATGAACATAT 420

QY 1593 TTGCTGGCTATACGGCGTAAATTAATCAATGGCTGTATGGACAGGCTATTCTTAACCGTCTGA 1652
DB 421 TTGCTGGCTATACGGCGTAAATTAATCAATGGCTGTATGGACAGGCTATTCTTAACCGTCTGA 480
QY 1653 CACCACTTTGTAGGCAATGCGCTTACGGTGCCTGCGCCAAAGTTTACCGCTCTATGATGACCT 1712
DB 481 CACCACTTTGTAGGCAATGCGCTTACGGTGCCTGCGCCAAAGTTTACCGCTCTATGATGACCT 540
QY 1713 ACCTGCTCTGAAGGAAGCAATCCAGAGGATTTGGAATATATACCAGAGGGGCTCTACAGAAATG 1772
DB 541 ACCTGCTCTGAAGGAAGCAATCCAGAGGATTTGGAATATATACCAGAGGGGCTCTACAGAAATG 600
QY 1773 GAGAATTCGTATTTAAAAATGGTGTCTGTCTACGTGGAATCTCACTGTCCACAAAC 1832
DB 601 GAGAATTCGTATTTAAAAATGGTGTCTGTCTACGTGGAATCTCACTGTCCACAAAC 660
QY 1833 CCCCATCAACTGAAAGTTCAAGCTCATCATCATCATCATCATCATCATCATCATCATCATCAT 1892
DB 661 CCCCATCAACTGAAAGTTCAAGCTCATCATCATCATCATCATCATCATCATCATCATCATCAT 720
QY 1893 CCACCTCCCAAGCACAAATATAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 1952
DB 721 CCACCTCCCAAGCACAAATATAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 780
QY 1953 ATACAACCCCTGATCAACAATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 1999
DB 781 ATACAACCCCTGATCAACAATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 827
RESULT 15
US-08-911-503-139
Sequence 139, Application US/08911503
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
NUMBER OF SEQUENCES: 713
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911.503
FILING DATE: 15-AUG-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,022
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50533
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 2172 base pairs
TYPE: nucleic acid

;
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-911-503-139

Query Match		40.5%;	Score 808.6;	DB 13;	Length 2172;
Best Local Similarity		99.4%;	Pred. No. 5.1e-218;		
Matches 822;		Conservative 0;	Mismatches 4;	Indels 1;	Gaps 1;
QY	1174	CGCCAGAGCTTTCCCTAAATGGTCTAGGAATCGACTACCCAAAGTATTTCAGTACTCAATGC	1233		
Db	1	CGCCAGAGCTTTCCCTAAATGGTCTAGGAATCGACTACCCAAAGTATTTCAGTACTCAATGC	60		
QY	1234	CATTTCAGTAACACAAACCGAATCAGACAAAAAATATGGAGCAAGTAGTGAAAAGATGGC	1293		
Db	61	CATTTCAGTAACACAAACCGAATCAGACAAAAAATATGGAGCAAGTAGTGAAAAGATGGC	120		
QY	1294	TGCTGCTTACGCTGCCCTTTGCAAAATGGTGGAACTTACTATATAACCAATGTATATCCATAA	1353		
Db	121	TGCTGCTTACGCTGCCCTTTGCAAAATGGTGGAACTTACTATATAACCAATGTATATCCATAA	180		
QY	1354	AGTCGCTTTTAGTGTGGAGTGAAAAGAGTCTCTAATGTCGGAACTCGTGCCCATGAA	1413		
Db	181	AGTCGCTTTTAGTGTGGAGTGAAAAGAGTCTCTAATGTCGGAACTCGTGCCCATGAA	240		
QY	1414	GGAAAGCAGAGCTATATGATGACCCGACATGATGAAAACAGTCTTGACTTATGGAACCTGG	1473		
Db	241	GGAAAGCAGAGCTATATGATGACCCGACATGATGAAAACAGTCTTGACTTATGGAACCTGG	300		
QY	1474	ACG-AAATGCCCTATCTTGGTCCCTCAGGCTGGTAAAAACAGGAACCTCTAACTATA	1532		
Db	301	ACGAAAATGCCCTATCTTGGTCCCTCAGGCTGGTAAAAACAGGAACCTCTAACTATA	360		
QY	1533	CAGAGGAGAAATGAAAACCAACATCAAGACCTCTCAATTTGTAGCACCTGATGAACCTAT	1592		
Db	361	CAGAGGAGAAATGAAAACCAACATCAAGACCTCTCAATTTGTAGCACCTGATGAACCTAT	420		
QY	1593	TTGCTGGCTATACGGGTAAATATTCAATGGCTGTATGGACAGGCTATTCTAACCCGTCTGA	1652		
Db	421	TTGCTGGCTATACGGGTAAATATTCAATGGCTGTATGGACAGGCTATTCTAACCCGTCTGA	480		
QY	1653	CACCACTGTAGGCAATGGCCTTACGGTCGCTGCCAAAGTTTACCGCTCTATGATGACCT	1712		
Db	481	CACCACTGTAGGCAATGGCCTTACGGTCGCTGCCAAAGTTTACCGCTCTATGATGACCT	540		
QY	1713	ACCTGTCTGAAGGAAATCCAGAGATTGGAAATATACAGAGGGGCTCTACAGAAATG	1772		
Db	541	ACCTGTCTGAAGGAAATCCAGAGATTGGAAATATACAGAGGGGCTCTACAGAAATG	600		
QY	1773	GAGAAATTCGTATTTAAAAATGGTCTCGTCTTACGTGGAACCTCACCTGCCACAACAAC	1832		
Db	601	GAGAAATTCGTATTTAAAAATGGTCTCGTCTTACGTGGAACCTCACCTGCCACAACAAC	660		
QY	1833	CCCCATCAACTGAAAGTTTCAAGCTCATCATAGATAGTTTCAACTTCAAGTCTAGCTCAA	1892		
Db	661	CCCCATCAACTGAAAGTTTCAAGCTCATCATAGATAGTTTCAACTTCAAGTCTAGCTCAA	720		
QY	1893	CCACTCCAGGCAATAATAGTACGACTACCAATCCCTAACAAATATACGACACATCAA	1952		
Db	721	CCACTCCAGGCAATAATAGTACGACTACCAATCCCTAACAAATATACGACACATCAA	780		
QY	1953	ATACACCCCTGATCAACAAAAATCAGAAATCCCTCAACGAGCAACCA	1999		
Db	781	ATACACCCCTGATCAACAAAAATCAGAAATCCCTCAACGAGCAACCA	827		

Search completed: June 13, 2002, 14:43:46
Job time: 7603 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 12:42:33 ; Search time 279.06 Seconds
(without alignments)
10068.144 Million cell updates/sec

Title: US-08-961-083-1

Perfect score: 1999
Sequence: 1 TAAATCTACGACATAA.....ATCCTCAACGACACCA 1999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 997330 seqs, 702755440 residues

Total number of hits satisfying chosen parameters: 1994660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US12_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	1999	100.0	1999	4	US-08-961-083-1		Sequence 1, Appl
2	117.2	5.9	9100	6	US-10-121-120-27		Sequence 27, Appl
3	59.2	3.0	2370	5	US-09-540-2098-3719		Sequence 3719, Ap
4	55.2	2.8	428	5	US-09-673-476-206		Sequence 206, App
5	51.8	2.6	363	5	US-09-673-476-626		Sequence 626, App
6	40.6	2.0	223	5	US-09-673-476-438		Sequence 438, App
7	40	2.0	610	6	US-10-027-632-5967		Sequence 5967, Ap
8	37.4	1.9	430	6	US-10-027-632-40693		Sequence 40693, A
9	37.4	1.9	242	5	US-09-789-189-1228		Sequence 1228, Ap
10	37.2	1.9	1611	7	US-60-360-039-46474		Sequence 46474, A
11	36.6	1.8	679	6	US-10-027-632-213202		Sequence 213202,
12	36.6	1.8	2409	7	US-60-360-039-46149		Sequence 46149, A
13	36.4	1.8	473	5	US-09-673-476-671		Sequence 671, App
14	36	1.8	689	6	US-10-027-632-148124		Sequence 148124,
15	35.8	1.8	348	5	US-09-673-476-99		Sequence 99, Appl
16	35.6	1.8	2373	5	US-09-540-2098-4136		Sequence 4136, Ap
17	35.4	1.8	8045	5	US-09-913-878A-1		Sequence 1, Appl
18	35	1.8	510	6	US-10-027-632-82144		Sequence 82144, A
19	35	1.8	510	6	US-10-027-632-309095		Sequence 309095,
20	35	1.8	132205	5	US-09-578-5198-25		Sequence 25, Appl
21	34.8	1.7	3435	1	PCR-US02-10624-12		Sequence 12, Appl
22	34.6	1.7	1247	6	US-10-027-632-254816		Sequence 254816,
23	34.6	1.7	1247	6	US-10-027-632-254817		Sequence 254817,
24	34.6	1.7	1247	6	US-10-027-632-254818		Sequence 254818,
25	34.6	1.7	1247	6	US-10-027-632-254819		Sequence 254819,
26	34.6	1.7	1247	6	US-10-027-632-254820		Sequence 254820,

27	34.6	1.7	2466	5	US-09-769-744A-121	Sequence 121, App
28	34.6	1.7	2486	6	US-10-104-047-1358	Sequence 1358, Ap
29	34.2	1.7	622	6	US-10-027-632-235624	Sequence 235624,
30	34.2	1.7	747	7	US-60-360-039-32748	Sequence 32748, A
31	34.2	1.7	2463	7	US-60-360-039-26589	Sequence 26589, A
32	34.2	1.7	14803	6	US-10-105-299-11424	Sequence 11424, A
33	34	1.7	585	6	US-10-027-632-219980	Sequence 219980,
34	34	1.7	653	6	US-10-123-155-438	Sequence 438, App
35	34	1.7	653	6	US-10-137-871-438	Sequence 438, App
36	34	1.7	653	6	US-10-141-761-438	Sequence 438, App
37	34	1.7	653	6	US-10-140-864-438	Sequence 438, App
38	34	1.7	653	6	US-10-140-923-438	Sequence 438, App
39	34	1.7	653	6	US-10-141-756-438	Sequence 438, App
40	34	1.7	653	6	US-10-141-759-438	Sequence 438, App
41	34	1.7	653	6	US-10-140-472-438	Sequence 438, App
42	34	1.7	653	6	US-10-140-805-438	Sequence 438, App
43	34	1.7	653	6	US-10-142-885-438	Sequence 438, App
44	34	1.7	653	6	US-10-146-731-438	Sequence 438, App
45	34	1.7	653	6	US-10-142-426-438	Sequence 438, App

ALIGNMENTS

RESULT 1
US-08-961-083-1
: Sequence 1, Application US/08961083
: GENERAL INFORMATION:
: APPLICANT: Choi et. al.
: TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
: NUMBER OF SEQUENCES: 452
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,083
: FILING DATE: 30-Oct-1997
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: <Unknown>
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1999 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-961-083-1

Query Match 100.0%; Score 1999; DB 4; Length 1999;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAATCTACGACATAAATCACTCAATTCGTTGGTCTGTAACGCCGCTCAA 60
|||||

Db 1 TAAATCTACGACATAAATAAATCAATCTATGCTGACTTGGGTCTTGAACGCCGCTCAA 60
 QY 61 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
 Db 61 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
 QY 121 TCGCTTCTTCGACACAGGGGATGATACCATCCGTATCCCTGGGAGCTTCTTCGCGAA 180
 Db 121 TCGCTTCTTCGACACAGGGGATGATACCATCCGTATCCCTGGGAGCTTCTTCGCGAA 180
 QY 181 TCTGCAACCAATTCCTCCAGGTGGATCACTCTCACCCACAGTTCGATTAAGTTGAC 240
 Db 181 TCTGCAACCAATTCCTCCAGGTGGATCACTCTCACCCACAGTTCGATTAAGTTGAC 240
 QY 241 TTACTTTTCAACTTCGACTTCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
 Db 241 TTACTTTTCAACTTCGACTTCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
 QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
 Db 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
 QY 361 GGTCTACATGCTTAATGGGAATATGGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 Db 361 GGTCTACATGCTTAATGGGAATATGGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 QY 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTAGCTTGGCTGGATGCAATCGCCAGCACC 480
 Db 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTAGCTTGGCTGGATGCAATCGCCAGCACC 480
 QY 481 AAACCAATATGACCCCTATTCACATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 Db 481 AAACCAATATGACCCCTATTCACATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGACAGATATGAGAAAGCAGTCAATAC 600
 Db 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGACAGATATGAGAAAGCAGTCAATAC 600
 QY 601 ACCAATTAAGTGGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 660
 Db 601 ACCAATTAAGTGGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 660
 QY 661 TAATTAAGTCAAGGAGTCAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 720
 Db 661 TAATTAAGTCAAGGAGTCAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 720
 QY 721 AACTGGGATGGATGCTACACAAATGCTACACCAAGAGCTCAAAACATCTGTGGGATAT 780
 Db 721 AACTGGGATGGATGCTACACAAATGCTACACCAAGAGCTCAAAACATCTGTGGGATAT 780
 QY 781 TTACATACAGAGATACGTTTGCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 840
 Db 781 TTACATACAGAGATACGTTTGCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 840
 QY 841 TGTGTATGTTTCAAGGTAAGTCAATTCGCCAGCTAGAGACAGCCATCAGTCAAGTAA 900
 Db 841 TGTGTATGTTTCAAGGTAAGTCAATTCGCCAGCTAGAGACAGCCATCAGTCAAGTAA 900
 QY 901 TGTGTCTTTCGGAATTAACCAAGCAGTACAAACCAACCGGAGTGGGATCAACTATGAA 960
 Db 901 TGTGTCTTTCGGAATTAACCAAGCAGTACAAACCAACCGGAGTGGGATCAACTATGAA 960
 QY 961 ACCGATACAGACTATGCTTCCCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT 1020
 Db 961 ACCGATACAGACTATGCTTCCCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT 1020
 QY 1021 CGTTACAGATGAGCCCTATAACTACCTGGGACAAATACCTCTGTTTATAACTGGGATAG 1080
 Db 1021 CGTTACAGATGAGCCCTATAACTACCTGGGACAAATACCTCTGTTTATAACTGGGATAG 1080
 QY 1081 GGGCTACTTTGGCAATCACTTCAATACCCCTGCAACCAATCGCGAAAGCTCCCGAG 1140
 Db 1081 GGGCTACTTTGGCAATCACTTCAATACCCCTGCAACCAATCGCGAAAGCTCCCGAG 1140

QY 1141 CGTGGAACTCTAAACAAGTGGGACTCAACCGCGCAAGACTTTCTCTAAATGCTTAGG 1200
 Db 1141 CGTGGAACTCTAAACAAGTGGGACTCAACCGCGCAAGACTTTCTCTAAATGCTTAGG 1200
 QY 1201 AATCGACTACCCCAAGTATTTCACTACTCAATGCCATTTCAAGTAACACACCAATCAGA 1260
 Db 1201 AATCGACTACCCCAAGTATTTCACTACTCAATGCCATTTCAAGTAACACACCAATCAGA 1260
 QY 1261 CAATAAATATGAGCAAGTAGTCAAAAGATGGCTGCTAGCTTAGCTTGGCTTTCGAATGG 1320
 Db 1261 CAATAAATATGAGCAAGTAGTCAAAAGATGGCTGCTAGCTTGGCTTTCGAATGG 1320
 QY 1321 TGAACCTTACTATAAACCAATGTATATCAATAAAGTCGCTTTAGTAGTGGGAGTGAAGA 1380
 Db 1321 TGAACCTTACTATAAACCAATGTATATCAATAAAGTCGCTTTAGTAGTGGGAGTGAAGA 1380
 QY 1381 AGAGTTCTCTAATGTGCGAAGTCTGCTGCAATGAAGGAAAGACAGCCTATATGATGACCA 1440
 Db 1381 AGAGTTCTCTAATGTGCGAAGTCTGCTGCAATGAAGGAAAGACAGCCTATATGATGACCA 1440
 QY 1441 CATGATGAAACAGTCTTGACTTATGGAATGGAACTGGACAAATGCCCTATCTTGGCTCCC 1500
 Db 1441 CATGATGAAACAGTCTTGACTTATGGAATGGAACTGGACAAATGCCCTATCTTGGCTCCC 1500
 QY 1501 TCAGGCTGTATAAACAGGAACCTCTAATCTATACAGACGAGGAAATTTGAAACCCACATCAA 1560
 Db 1501 TCAGGCTGTATAAACAGGAACCTCTAATCTATACAGACGAGGAAATTTGAAACCCACATCAA 1560
 QY 1561 GACCTCTCAATTTGTAGTACCTGATGAATTTTGTGCTATACGCGTAAATATTTCAAT 1620
 Db 1561 GACCTCTCAATTTGTAGTACCTGATGAATTTTGTGCTATACGCGTAAATATTTCAAT 1620
 QY 1621 GCTGTATGACAGAGGCTATTTAAACCGTCTGACACCACTTTGAGGCAATGGCCTTACGGT 1680
 Db 1621 GCTGTATGACAGAGGCTATTTAAACCGTCTGACACCACTTTGAGGCAATGGCCTTACGGT 1680
 QY 1681 CGCTGCAAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAGGAAAGCAATCCAGAAGA 1740
 Db 1681 CGCTGCAAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAGGAAAGCAATCCAGAAGA 1740
 QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 1800
 Db 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 1800
 QY 1801 TTCTACGTGAACTACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
 Db 1801 TTCTACGTGAACTACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
 QY 1861 ATCAGATGTTCAACTTCAAGTCTACAGTCTAGCTCAACCCACTCCAGCACAATAATAGTAGCAG 1920
 Db 1861 ATCAGATGTTCAACTTCAAGTCTACAGTCTAGCTCAACCCACTCCAGCACAATAATAGTAGCAG 1920
 QY 1921 TACCATCTTCAATAATACGCAACAATCAATAACACCCCTGATCAACAAATCAGAA 1980
 Db 1921 TACCATCTTCAATAATACGCAACAATCAATAACACCCCTGATCAACAAATCAGAA 1980
 QY 1981 TCCTCAACAGCACAACCA 1999
 Db 1981 TCCTCAACAGCACAACCA 1999

RESULT 2

US-10-121-120-27/c

; Sequence 27, Application US/10121120

; GENERAL INFORMATION:

; APPLICANT: Bergeron, Michel G.

; APPLICANT: Ouellette, Marc

; APPLICANT: Roy, Paul H.

; TITLE OF INVENTION: Specific and Universal Probes and Amplification

; TITLE OF INVENTION: Primers

; TITLE OF INVENTION: to Rapidly Detect and Identify Common Bacterial Pathogens and Antibiotic Resistance Genes from Clinical


```

; TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro
; FILE REFERENCE: 12287.31
; CURRENT APPLICATION NUMBER: US/10/121,120
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/452,599
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: 08/304,732
; PRIOR FILING DATE: 1994-09-12
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 9100
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; US-10-121-120-27

```

Query Match	5.9%	Score 117.2	DB 6	Length 9100
Best Local Similarity	50.3%	Pred. No. 1.5e-22		
Matches 349	Conservative 0	Mismatches 33	Indels 12	Gaps 2
QY 22	TCAACTCATTGCTGACATTGGTTCCTGACGCCGCGTCAATGCCCAAGCTAATGATATCC 81			
Db 3498	TAAATTAATTGGCGAAGTGGGTGAGCAACGCCGATTCCAGTGAATTAGCCGATGTGCC 3439			
QY 82	CACAGATTGGTTAAGCAATCGTTCTATCGAAGACCATCGCTCTCTCGACACACAGGG 141			
Db 3438	ACACGCTTAATTGACGCATTTTAGCGCGGAAGACAGTCGTTTTCAGATCATCACGG 3379			
QY 142	GATTGATACCATCGGTATCTCGGGAGCTTTCTTCGGCAATCTGCAAA---GCAATTCCCT 198			
Db 3378	ATTAGACCCATTGCGCATTTGCCGTGCATTTGTCGCACTGAGTAATGCGCGTGCATC 3319			
QY 199	CCAGGTTGGATCAACTCTCACCCAAACAGTTGATTAAAGTTGACTTACATTTTCAACTTCGAC 258			
Db 33:8	ACAAGCGCAAGTACGATTACTCAACAATTAGCCGTAACCTTTTCTCTA-----AC 3268			
QY 259	TTCCGACACAGACTATTTCCTGAAGGCTCAGGAAGCTTGGTTAGCGATTCAAGTTAGAACA 318			
Db 3267	CTCAGAAAAACCATTTTCGTAAGCTCGTGAAGCCGCTGCTCGCGTAGAAATCGAAAA 3208			
QY 319	AAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAAATAAGTCTACATGTCTAATGG 378			
Db 3207	TACTCTCAACAAACAGAAATATTAGAGCTTTATTAAACAAAAATCTTTTAGGCTATCG 3148			
QY 379	GNACTATGGATGCAGACAGCAGCTCAAACTACTACTATGTTAAAGACCTCAATATTTTAAG 438			
Db 3147	TTCTATATGGTGTGACGGCGCAGCACAAACCTATTTCGGTAANATCATTTGAATGAATTGAC 3088			
QY 439	TTTACCTCAGTTAGCCCTTGCTGGTGGAAATGCCCTCAGGCACCAACCAATATGACCCCTA 498			
Db 3087	CTTATCGGAATGCGGATATTGCTGGTTTACCTTAAGCACCTTCAACAATGAACCCGCT 3028			
QY 499	TTCAATCCAGACGACGCCCAAGACCCGCAAACTTGGTCTTATCTGAAATGAAAAATCA 558			
Db 3027	TTATTCTTTAAAAGGTTTCAGAAAGACCCGCAATGTGGTGTGCTAAGCCGTATTAGATGA 2968			
QY 559	AGGCTACATCTCTGCTCAACAGATATGAAAGCAGTCAATACCAATTACTGATGGACT 618			
Db 2967	AAATACATCAGCAAAAGAAATATGATGTGCATTGAAAGACGCCGATTTGGCGGAGCTA 2908			
QY 619	ACAAGTCTCAATCAACGAAGTATTACCCTGCTTACATGGATTAATTAACCTCAAGGAAGT 678			
Db 2907	TCAGGGCGCAAAATTTGAATTTCCAGCCGATATTGTCACATGAATGGTGGCTCAAGAAAT 2848			
QY 679	CATCAATCAAGTTGAAGAAGAAACAGGCTATAAC 712			
Db 2847	GGTGCCTGCTTTTGGCGAAGAAATGCTTCACCC 2814			

RESULT 3
US-09-540-209B-3719
: Sequence 3719, Application US/09540209B

```

; GENERAL INFORMATION:
; APPLICANT: Gary L. Bregon
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 3719
; LENGTH: 2370
; TYPE: DNA
; ORGANISM: B.fragilis
; US-09-540-209B-3719

```

Query Match	3.0%;	Score 59.2;	DB 5;	Length 2370;
Best local Similarity	46.9%;	Pred. No. 2.2e-06;		
Matches 261;	Conservative 0;	Mismatches 283;	Indels 12;	Gaps 2;
Qy	42	GTTCGAAAGCGCGTCAATCCCAAGCTAATGATATATCCACAGATTGTTGAAGCAA	101	
Db	233	gtaaggaaaaatcggtgtatacacaccttataatgaacttccccaacattgtcaatgat	292	
Qy	102	TCGTTTCTATCGAAGACCATCTCTTCGACCACAGGGGATTGATACCATCCGTATCC	161	
Db	293	tgatcgccagaggaagcttcgctttaccgaaattcgggtatcgatgccaaagcgtga	352	
Qy	162	TGGGAGCTTCTT-----GGCRAATGCAAGCAATTCCTCCAGAGTGGATCAATC	215	
Db	353	tacgtgtgtgtaaaagcgtgattgctgtatgcagaaaaatgcaggtggagcgatcac	412	
Qy	216	TCACCCAAACAGTGTATTA-----GTGTGACTTACTTTTCAACTTCGACTTC	269	
Db	413	ttcacacaagactcgccaaagcaattgtttacg9acgaagtgtccagaataacgctgcgc	472	
Qy	270	CTATTTCCTGAAGGCTCAGGAAGCTTGGTTAGCGGATTACGTTAGAACAAAAAGCA	329	
Db	473	gcctgtttcagaagccgatagagtggtggtgattggtcgtaaaactggaacgttattataca	532	
Qy	330	AGCAAGAAATTCACCTACTATATAATAGGTCTACATGCTTAATGGGAACATATGGAA	389	
Db	533	aggaagaatttgatgatgtatctcaataaattgacttctgtaataatgcagtaggaa	592	
Qy	390	TGCAGACGACGCTCAAACTPACTATGATGTAAGACCTCAATAATTAAAGTTTACCT	449	
Db	593	ttaaaacggttcataatccctatttcggatcggaacccaagaatctgaaatagaacaag	652	
Qy	450	TAGCCTTGCTGGGTGGGAATGCGCTCAGGCACCAACCAATATGACCCCTATTTCACAT	509	
Db	653	ctgctacgctgatcggtatgtgtaaaaactcttcgctttaaactccggtgcgcttcaacg	712	
Qy	510	AAGCAGCCCAAGACCGCGGAAACTTGGTCTTATCTGTAATCAAAATCAAGGCTACATCT	569	
Db	713	agcgttcgcgcgacgcggaataacagtctagatcctaagatgagaaagccggatatacca	772	
Qy	570	CTGCTGAACAGTATCA	585	
Db	773	cagcgaagactgtga	788	

RESULT 4
US-09-673-476-206
; Sequence 206, Application US/09673476
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; APPLICANT: BUCHRIESER-BROSCH, ROLAND
; APPLICANT: GORDON, STEPHEN
; APPLICANT: BILLAULT, ALAIN
; TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
; TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
; TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF
; TITLE OF INVENTION: MYCOBACTERIA.
; FILE REFERENCE: 05394. 0011-000000

QV 401 GCTCAAAC TACTATGGTAAAGACCTCAATAATTTAAGTTTACCTCAGTTAGCCCTTGCTG 460


```
Db 126 gcgcaaacgctnctcgcatcaacgcgtcganctgaattggcagcaagcgcgctgctg 185
QY 461 GCTGGAATGCTCAGGCAACCAACCAATATGACCCCTA 498
Db 186 gcgcgcatggtgcaatcaccagcagcgtcaaccgta 223

RESULT 7
US-10-027-632-5967/c
; Sequence 5967, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5967
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-5967

Query Match 2.0%; Score 40; DB 6; Length 610;
Best Local Similarity 55.1%; Pred. No. 0.34; Indels 0; Gaps 0;
Matches 76; Conservative 1; Mismatches 61;

QY 1845 AAGATTCAAGCTCATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCCCAAGCA 1904
Db 410 AGAGYGAACTCCATCTCAATACTACTACTACTACTACTACTACTACTACTACTACTA 351
QY 1905 CAAATAAATAGTAGGACTCAATCTTAACAATAATACGCAACAATCAATACAAACCCCTG 1964
Db 350 CTACTACTACTACTATAATAATAATAATAATAATGAGAGCTAAACATATTTACCTG 291
QY 1965 ATCACAAAATCAGATC 1982
Db 290 TAGGCGCACTATCTCAATC 273

RESULT 8
US-10-027-632-40693/c
; Sequence 40693, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
```

```
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40693
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(430)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-40693

Query Match 1.9%; Score 38; DB 6; Length 430;
Best Local Similarity 60.8%; Pred. No. 1.1; Indels 0; Gaps 0;
Matches 62; Conservative 0; Mismatches 40;

QY 319 AAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAAAATAGGTCTCATGTCTATGG 378
Db 392 AAAAGCAAGTGGGAAAAAACAACAACCAACAATAAAATGAGAGCTGAATTTTAAACGG 333
QY 379 GAACATATGAATCCAGACAGCAGCTCAAAACTACTATGTGTA 420
Db 332 AAGCAATGGAATTCAGAAGCAATAAGATAATATTTTAAAA 291

RESULT 9
US-09-789-189-1228
; Sequence 1228, Application US/09789189
; GENERAL INFORMATION:
; APPLICANT: Lelias, Jean-Michel
; TITLE OF INVENTION: Human Polynucleotides and Polypeptides
; FILE REFERENCE: 25436/1720
; CURRENT APPLICATION NUMBER: US/09/789,189
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/183452
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 2005
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1228
; LENGTH: 242
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(242)
; OTHER INFORMATION: "n" at position 33 refers to any of G, A, T or C
US-09-789-189-1228

Query Match 1.9%; Score 37.4; DB 5; Length 242;
Best Local Similarity 54.4%; Pred. No. 1.3; Indels 0; Gaps 0;
Matches 74; Conservative 0; Mismatches 62;

QY 308 CAGTTAGACAAAAGCAAGCAAGAAATCTTGACCTACTATATAAAATAGGTCTAC 367
Db 32 cnatcaaaaaaaaaaaaaaaaaaaaaaacaaattatcatatacaaacataataaata 91
QY 368 ATGCTTAATGGGAACATATGAATGCAGACAGCAGCTCAAAACTACTATGTGTAAGACCTC 427
Db 92 caatccatctacaacataaaaaataaataaacttaaacctcctattataaacatctt 151
QY 428 AATAATTTAAGTTTAC 443
```


Db 152 tataataaactaac 167

RESULT 10

US-60-360-039-46474

; Sequence 46474, Application US/603600039

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Chen, Xianfeng

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)A

; CURRENT APPLICATION NUMBER: US/60/360,039

; CURRENT FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 46474

; LENGTH: 1611

; TYPE: DNA

; ORGANISM: Schizosaccharomyces pombe

US-60-360-039-46474

Query Match

; Sequence 37.2; DB 7; Length 1611;

; Best Local Similarity 53.4%; Pred. No. 3;

Matches 78; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 746 GTAGACCAAGAGCTCAAAAACATCTGTGGATATTTACAATACAGACGAATACGTTGCC 805

DB 829 ggagatgaagtggtgtatccttgcctattttatattcaagaatacaaaaacgttgg 888

QY 806 TATCCAGACGAATGTAAGTCGCTTCTACCATGTTGATGTTCTACGGTAAAGTC 865

DB 889 aagccaattgatgattttacgttttgcactactgtctctagtctatggttaaagtc 948

QY 866 ATTGCCAGCTAGGACGCGCATCA 891

DB 949 atggccaaagtagaaggattccatca 974

RESULT 11

US-10-027-632-213202/c

; Sequence 213202, Application US/10027632

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; TITLE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 213202

; LENGTH: 679

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-213202

Query Match 1.8%; Score 36.6; DB 6; Length 679;

Best Local Similarity 49.2%; Pred. No. 3.2;

Matches 96; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 683 AATCAAGTTGAAGAGAAACAGGCTATAACCTACTCACAACCTGGGATGGTCTACACA 742

DB 568 AAACCTAAGAAATAGAAATCAGATTTTAAATATGTTAGGTTCAAACTCAAAAACA 509

QY 743 AATGTAGACCAAGAGCTCAAAAACATCTGTGGATATTTACAATACAGACGAATACGTT 802

DB 508 AAGAGAAATACAGATCAACACATACAAATCACTATAAAATTTATTAAAAATTAAGAAGTT 449

QY 803 GCCTATCCAGAGATGAATTCGAAGTCGCTTCTACCATGTTGATGTTTCTACACGCTAAA 862

DB 448 CCTTTTCTCCAAGATCTTTTAAAGTTGTTTTGGTTTATAGATACTACTTCTGACTTTCAA 389

QY 863 GTCAATTGCCAGCTA 877

DB 388 ACAATTTCAAACTA 374

RESULT 12

US-60-360-039-46149

; Sequence 46149, Application US/603600039

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Chen, Xianfeng

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)A

; CURRENT APPLICATION NUMBER: US/60/360,039

; CURRENT FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 46149

; LENGTH: 2409

; TYPE: DNA

; ORGANISM: Saccharomyces cerevisiae

US-60-360-039-46149

Query Match

1.8%; Score 36.6; DB 7; Length 2409;

Best Local Similarity 52.3%; Pred. No. 5.2;

Matches 81; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 673 GGAAGTCATCAATCAAGTTGAAGAGAAACAGGCTATAACCTACTCACAACCTGGGATGGA 732

DB 243 ggaagtaataaacacagaagaagaagaagaacatcaacaaaagcggaatga 302

QY 733 TGTCTACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATATTTACAATACAGA 792

DB 303 tgatgacgatgatgaataatgaagagggaagaggagatgatgacgatga 362

QY 793 CGAATACGTTGCCTATCCAGACGATGAATTCGAAG 827

DB 363 tgatgacgacgatgatgatgaagaagaagaagaag 397

RESULT 13

US-09-673-476-671/C

; Sequence 671, Application US/09673476

; GENERAL INFORMATION:

; APPLICANT: COLE, STEWART

; APPLICANT: BUCHRIESER-BROSCH, ROLAND

; APPLICANT: GORDON, STEPHEN

; APPLICANT: BILLAULT, ALAIN

; TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST

; TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED

; TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF

; TITLE OF INVENTION: MYCOBACTERIA.
; FILE REFERENCE: 05394.0011-00000
; CURRENT APPLICATION NUMBER: US/09/673,476
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: PCT/IB99/00740
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 09/060,756
; PRIOR FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: Patentin Ver. 2.2
; SEQ ID NO 671
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (147)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (247)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (380)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (457)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (468)
; OTHER INFORMATION: a, t, c or g
; US-09-673-476-671

Query Match 1.8%; Score 36.4; DB 5; Length 473;
Best Local Similarity 51.6%; Pred. No. 3.2;
Matches 82; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 5 ATCTAGCACAATAAATCAACTCATCTGGTTCGAGTTCGAGCGCGGTCATGCC 64
Db 183 ATGGTCGACGGAAGGCAACACGCGTGGCTGACTCGAGCGCGGTTTCGAGGTG 124
QY 65 CAAGCTTAATGATATCCCAAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCATCGC 124
Db 123 CCTCGGACAAGATCCCAACAGATGAAGCTGGCGATCGTCTCGATTGAAGATAACGGG 64
QY 125 TTCTTCGACACAGGGGATTTGATACCATCCGTPATCCTG 163
Db 63 TTGCGGACACAGCGCGTGGACTGGAAGGGCACCCCTG 25

RESULT 14
US-10-027-632-148124/c
; Sequence 148124, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148124
; LENGTH: 689
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-148124

Query Match 1.8%; Score 36; DB 6; Length 689;
Best Local Similarity 51.2%; Pred. No. 4.8;
Matches 84; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 1832 CCCCCATCAACTGAAAGTTCAAGCTCATCATCAGATAGTTCAACTTCACAGTCTAGTCA 1891
Db 360 CCCCCCTCCAGAGGATAATACTCCCTCTAGAGATAATACTCCTCTCCAGAGATAAT 301
QY 1892 ACCACTCCAAGCACAAATAATAGTACGACTACCAATCCTTAACAATAATACGCAACATCA 1951
Db 300 ACTCCTCTCCAGAGATAATACTCCTCTCCAGAGATAATACTCCTCTCCAGAA 241
QY 1952 AATAGAACCCCTGATCAACAAATCAGATCCTCAACACGACACA 1995
Db 240 GATAATACTCCTCTCTAGAGATAATACTCCTCTCCAGAGAA 197

RESULT 15
US-09-673-476-99/c
; Sequence 99, Application US/09673476
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; APPLICANT: BUCHRIER-BROSCH, ROLAND
; APPLICANT: GORDON, STEPHEN
; APPLICANT: BILLAULT, ALAIN
; TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
; FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
; FILE REFERENCE: 05394.0011-00000
; CURRENT APPLICATION NUMBER: US/09/673,476
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: PCT/IB99/00740
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 09/060,756
; PRIOR FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: Patentin Ver. 2.2
; SEQ ID NO 99
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; US-09-673-476-99

Query Match 1.8%; Score 35.8; DB 5; Length 348;
Best Local Similarity 51.6%; Pred. No. 4.3;
Matches 82; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 5 ATCTAGCACAATAAATCAACTCATCTGGTTCGAGTTCGAGCGCGGTCATGCC 64
Db 321 ATGGTCGACGGAAGGCAACACGATCGCGTGGCTGACTCGAGCGCGGTTTCGAGGTG 262
QY 65 CAAGCTTAATGATATCCCAAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCATCGC 124
Db 261 CCTCGGACAAGATCCCAACAGATGAAGCTGGCGATCGTCTCGATTGAAGATAACGGG 202
QY 125 TTCTTCGACACAGGGGATTTGATACCATCCGTPATCCTG 163
Db 63 TTGCGGACACAGCGCGTGGACTGGAAGGGCACCCCTG 25

Db 201 TTGCGCGACACAGCGGCGTGGACTGGAAGGGCACCCCTG 163

Search completed: June 13, 2002, 14:47:52
Job time: 7519 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 08:39:33 ; Search time 3635.99 seconds
(without alignments)
11505.022 Million cell updates/sec

Title: US-08-961-083-1

Perfect score: 1999

Sequence: 1 TAAATCTACGACAAATRAAA.....ATCCTCAACGACCAACCA 1999

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 15

Total number of hits satisfying chosen parameters: 37854

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hgt:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_hgt_hum:*

31: em_hgt_inv:*

32: em_hgt_other:*

33: em_hgtg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	1999	100.0	1999	6	AR120238	Sequence
2	1999	100.0	10711	6	BD003825	Polynucle
3	1999	100.0	11334	1	AE007349	Streptoco
4	1795	89.8	2160	1	AF210745	Streptoco
5	1591	79.6	2157	1	SPPONAS	S.pneumonia
6	1489	74.5	2160	1	AF210746	Streptoco
7	1438	71.9	2157	1	SPPONAS	S.pneumonia
8	1438	71.9	2160	6	AX111315	Sequence
9	1438	71.9	3378	1	STRPONA	M90527 Streptococ
10	1438	71.9	10313	1	AE008414	Streptoco
11	1198	59.9	1249	6	AX111327	Sequence
12	889	44.5	1195	1	AB006874	Streptoco
13	835	41.8	1195	1	AB006868	Streptoco
14	796	39.8	960	6	AR084885	Sequence
15	774	38.7	1826	1	SPPONAS	S.pneumonia
16	705	35.3	2154	1	AF210748	Streptoco
17	671	33.6	930	1	AF046233	Streptoco
18	620	31.0	930	1	AF046235	Streptoco
19	620	31.0	930	1	AF046237	Streptoco
20	620	31.0	930	6	AX111320	Sequence
21	590	29.5	1199	6	AX110281	Sequence
22	567	28.4	1229	6	AX110284	Sequence
23	551	27.6	930	1	AF046236	Streptoco
24	461	23.1	930	1	AF046234	Streptoco
25	436	21.8	2160	1	AF210747	Streptoco
26	436	21.8	30457	2	SPNEU1914	Streptoco
27	432	21.6	2160	6	AX111314	Sequence
28	396	19.8	2172	1	SPCS1091A	S.pneumonia
29	396	19.8	2172	1	SPCS1111A	S.pneumonia
30	396	19.8	2172	1	SPN403975	Streptoco
31	387	19.4	489	1	SPN403978	Streptoco
32	387	19.4	2157	1	SPPONAS	S.pneumonia
33	318	15.9	2157	1	SPPONAS	S.pneumonia
34	318	15.9	2157	1	SPPONAS	S.pneumonia
35	318	15.9	2157	1	SPPONAS	S.pneumonia
36	318	15.9	2157	1	SPPONAS	S.pneumonia
37	318	15.9	2157	6	AX111407	Sequence
38	318	15.9	2160	1	AF139883	Streptoco
39	318	15.9	2160	1	AF139884	Streptoco
40	318	15.9	2160	1	AF139885	Streptoco
41	318	15.9	2160	1	AF139886	Streptoco
42	318	15.9	2160	1	AF139887	Streptoco
43	318	15.9	2160	1	AF139888	Streptoco
44	318	15.9	2160	1	AF139889	Streptoco
45	318	15.9	2160	1	AF139890	Streptoco
46	318	15.9	2160	1	AF159448	Streptoco
47	318	15.9	2160	1	AF210749	Streptoco
48	318	15.9	2160	1	AF210750	Streptoco
49	318	15.9	2160	1	AF210751	Streptoco
50	318	15.9	2160	6	AX110436	Sequence
51	318	15.9	2160	6	AX111316	Sequence
52	298	14.9	1195	1	AB006869	Streptoco
53	298	14.9	1195	1	AB006872	Streptoco
54	298	14.9	1195	1	AB006873	Streptoco
55	245	12.3	1195	1	AB006879	Streptoco
56	228	11.4	1195	1	AB006877	Streptoco
57	228	11.4	1195	6	AX111319	Sequence
58	224	11.2	567	1	SPPBP1AA	S.pneumonia
59	209	10.5	489	1	SPN403977	Streptoco
60	205	10.3	1195	1	AB006870	Streptoco
61	205	10.3	1195	1	AB006878	Streptoco
62	203	10.2	930	1	AF046232	Streptoco
63	197	9.9	930	1	AF046231	Streptoco
64	178	8.9	1212	6	AX111323	Sequence
65	178	8.9	1222	6	AX110283	Sequence
66	175	8.8	1220	6	AX110280	Sequence
67	159	8.0	1216	6	AX111329	Sequence
68	159	8.0	2157	1	SPPONAS	S.pneumonia
69	153	7.7	1195	1	AB006876	Streptoco
70	148	7.4	1195	1	AB006871	Streptoco
71	131	6.6	489	1	SPN403979	Streptoco
72	129	6.5	1201	6	AX110279	Sequence
73	127	6.4	1212	6	AX110272	Sequence

74	127	6.4	1223	6	AX110275	AX110275 Sequence	147	20	1.0	113688	9	AC005301	AC005301 Homo sapi
75	126	6.3	1211	6	AX110282	AX110282 Sequence	c 148	20	1.0	118969	8	H0212B02	AL4242007 Oryza sat
76	126	6.3	1218	6	AX110274	AX110274 Sequence	c 149	20	1.0	119770	8	AC008697	AC008697 Homo sapi
77	122	6.1	1213	6	AX110273	AX110273 Sequence	150	20	1.0	124823	9	AC007064	AC007064 Homo sapi
78	122	6.1	1214	6	AX110276	AX110276 Sequence	151	20	1.0	139634	2	AP004223	AP004223 Oryza sat
79	121	6.1	1207	6	AX110278	AX110278 Sequence	152	20	1.0	141371	2	AC084087	AC084087 Homo sapi
80	120	6.1	1212	6	AX110271	AX110271 Sequence	153	20	1.0	145684	2	AC016588	AC016588 Homo sapi
81	116	5.8	489	1	SPN403980	AJ403980 Streptococ	c 154	20	1.0	146120	9	AC024942	AC024942 Homo sapi
82	115	5.8	1242	6	AX111324	AX111324 Sequence	c 155	20	1.0	148183	2	AP004672	AP004672 Oryza sat
83	110	5.5	1223	6	AX111027	AX111027 Sequence	c 156	20	1.0	151024	2	AC016855	AC016855 Homo sapi
84	104	5.2	1225	6	AX111325	AX111325 Sequence	c 157	20	1.0	158344	10	AL133401	AL133401 Mouse DNA
85	97	4.9	930	1	AF046230	AF046230 Streptococ	158	20	1.0	158467	9	AC011314	AC011314 Homo sapi
86	97	4.9	930	6	AX111317	AX111317 Sequence	159	20	1.0	161735	9	AC068633	AC068633 Homo sapi
87	90	4.5	782	6	AX111331	AX111331 Sequence	160	20	1.0	161788	9	AC009314	AC009314 Homo sapi
88	90	4.5	813	6	AX110915	AX110915 Sequence	161	20	1.0	165725	2	OSJN00105	AL060692 Oryza sat
89	77	3.9	489	1	SPN403976	AJ403976 Streptococ	162	20	1.0	166204	2	AC026763	AC026763 Homo sapi
90	73	3.7	1225	6	AX110285	AX110285 Sequence	c 163	20	1.0	167063	2	AL353758	AL353758 Homo sapi
91	44	2.2	489	1	SPN403974	AJ403974 Streptococ	c 164	20	1.0	168720	2	AC078781	AC078781 Homo sapi
92	44	2.2	489	1	SPN403981	AJ403981 Streptococ	165	20	1.0	169340	9	AP001148	AP001148 Homo sapi
93	44	2.2	930	1	AF046238	AF046238 Streptococ	c 166	20	1.0	170232	9	AL135903	AL135903 Human DNA
94	44	2.2	930	1	AF387163	AF387163 Streptococ	c 167	20	1.0	170751	2	HSBA536C5	AL121987 Homo sapi
95	44	2.2	930	6	AX111318	AX111318 Sequence	168	20	1.0	175933	2	AC096595	AC096595 Rattus no
96	44	2.2	1582	1	SMAJ2290	AJ002290 Streptococ	169	20	1.0	176236	2	AC026344	AC026344 Homo sapi
97	34	1.7	2181	1	SMI295856	AJ295856 Streptococ	c 170	20	1.0	177564	9	AP002799	AP002799 Homo sapi
98	29	1.5	1201	1	AB006875	AB006875 Streptococ	171	20	1.0	180319	2	AC016918	AC016918 Homo sapi
99	26	1.3	930	1	AY072702	AY072702 Streptococ	c 172	20	1.0	187985	2	AL611932	AL611932 Homo sapi
100	25	1.3	25	6	AX110484	AX110484 Sequence	c 173	20	1.0	188078	9	AC008045	AC008045 Homo sapi
101	23	1.2	33	6	AR120351	AR120351 Sequence	c 174	20	1.0	194334	2	AC046140	AC046140 Homo sapi
102	23	1.2	2528	1	STRP0NNA	M90528 Streptococ	c 175	20	1.0	194674	2	AC021168	AC021168 Homo sapi
103	22	1.1	22	6	AX110459	AX110459 Sequence	176	20	1.0	195349	2	AC006705	AC006705 Caenorhab
104	22	1.1	4534	1	AX111282	AX111282 Sequence	177	20	1.0	197110	9	AC104306	AC104306 Homo sapi
105	22	1.1	4534	1	EFA302065	AJ302065 Enterococ	c 178	20	1.0	197110	17	AC010742	AC010742 Homo sapi
106	22	1.1	189452	2	AL663110	AL663110 Mus muscu	c 179	20	1.0	201064	2	AL627236	AL627236 Mus muscu
107	21	1.1	138317	9	AP003481	AP003481 Homo sapi	c 180	20	1.0	201215	9	AC019084	AC019084 Homo sapi
108	21	1.1	152816	2	AC079739	AC079739 Homo sapi	c 181	20	1.0	214187	9	AC008703	AC008703 Homo sapi
109	21	1.1	162608	9	AC009478	AC009478 Homo sapi	c 182	20	1.0	234270	2	AL627342	AL627342 Mus muscu
110	21	1.1	167322	2	AC016788	AC016788 Homo sapi	183	19	1.0	19	6	A83833	A83833 Sequence 11
111	21	1.1	184767	2	AC084363	AC084363 Homo sapi	c 184	19	1.0	19	6	AY37187	AY37187 Arabidops
112	21	1.1	302387	2	AC026348	AC026348 Homo sapi	c 185	19	1.0	382	8	AY37187	AY37187 Arabidops
113	20	1.0	20	6	AX110392	AX110392 Sequence	186	19	1.0	465	11	G49014	G49014 SHGC-78177
114	20	1.0	20	6	AX111291	AX111291 Sequence	c 187	19	1.0	864	33	AC071669	AC071669 Giardia i
115	20	1.0	20	6	E14593	E14593 PCR primer	c 188	19	1.0	914	33	AC039883	AC039883 Giardia i
116	20	1.0	20	6	E14594	E14594 PCR primer	c 189	19	1.0	915	33	AC089821	AC089821 Giardia i
117	20	1.0	25	6	AX110485	AX110485 Sequence	c 190	19	1.0	935	33	AC046483	AC046483 Giardia i
118	20	1.0	550	6	AX111543	AX111543 Sequence	c 191	19	1.0	1050	33	AC059516	AC059516 Giardia i
119	20	1.0	554	6	AX111326	AX111326 Sequence	192	19	1.0	1065	1	HEA28KDM	M59804 Haemophilus
120	20	1.0	680	6	A97193	A97193 Sequence 39	193	19	1.0	1065	33	AC039210	AC039210 Giardia i
121	20	1.0	1290	9	BC010387	BC010387 Homo sapi	c 194	19	1.0	1126	5	FSAPOB	M15892 Petromyzon
122	20	1.0	1620	14	AF425763	AF425763 Human cal	c 195	19	1.0	1922	8	AF201456	AF201456 Arabidops
123	20	1.0	1620	14	AF425764	AF425764 Human cal	196	19	1.0	2193	6	AR020783	AR020783 Sequence
124	20	1.0	1620	14	AF425766	AF425766 Human cal	197	19	1.0	2193	6	AR020784	AR020784 Sequence
125	20	1.0	1620	14	AF427115	AF427115 Human cal	c 198	19	1.0	2195	9	AF281970	AF281970 Homo sapi
126	20	1.0	1620	14	AF427116	AF427116 Human cal	199	19	1.0	2196	1	AF101780	AF101780 Streptoco
127	20	1.0	1620	14	AF427121	AF427121 Human cal	200	19	1.0	2203	10	MM085259	U85259 Mus musculo
128	20	1.0	1909	5	AB024280	AB024280 Pelodiscu	c 201	19	1.0	2208	1	SPAJ2292	AJ002292 Streptoco
129	20	1.0	2184	1	SMI295852	AJ295852 Streptoco	c 202	19	1.0	6906	6	AX344602	AX344602 Sequence
130	20	1.0	3472	10	MUSNEDD1	D10712 Mus musculo	203	19	1.0	7356	1	AF288082	AF288082 Photobab
131	20	1.0	4156	9	AF040250	AF040250 Homo sapi	c 204	19	1.0	10146	1	U32699	U32699 Haemophilus
132	20	1.0	4508	9	HSPOGAGNA	Z22642 H.sapiens P	c 205	19	1.0	10856	1	AE008546	AE008546 Streptoco
133	20	1.0	4508	9	HUMPOGA	L14922 Homo sapien	c 206	19	1.0	11486	1	U32744	U32744 Haemophilus
134	20	1.0	4952	9	HUMRFC	L23320 Human repli	c 207	19	1.0	11486	1	U32744	U32744 Haemophilus
135	20	1.0	5124	9	AE0028958	AB028958 Homo sapi	c 208	19	1.0	18662	3	AF381613	AF381613 Ancylost
136	20	1.0	10061	1	AE004314	AE004314 Vibrio ch	c 209	19	1.0	35447	2	AC007181	AC007181 Homo sapi
137	20	1.0	29101	3	CE2K1058	Z35604 Caenorhabdi	c 210	19	1.0	38583	2	AC017825	AC017825 Drosophill
138	20	1.0	39816	2	AC110404	AC110404 Rattus no	c 211	19	1.0	38868	3	U28941	U28941 Caenorhabdi
139	20	1.0	46555	3	CEY57G11A	Z99279 Caenorhabdi	212	19	1.0	43361	2	AC094627	AC094627 Rattus no
140	20	1.0	48060	9	AL139798	AL139798 Human DNA	213	19	1.0	45980	2	U82206	U82206 Homo sapien
141	20	1.0	57233	2	AC100137	AC100137 Mus muscu	214	19	1.0	58442	2	AC109278	AC109278 Mus muscu
142	20	1.0	66924	8	NCEB9J10	AL356324 Neurospor	215	19	1.0	60543	2	AC100307	AC100307 Mus muscu
143	20	1.0	70564	9	HSBJ346G2	AL049546 Human DNA	216	19	1.0	62355	3	AC084472	AC084472 Caenorhab
144	20	1.0	84023	9	AL451131	AL451131 Human DNA	217	19	1.0	67698	2	AC105979	AC105979 Mus muscu
145	20	1.0	92348	2	AC094515	AC094515 Rattus no	c 218	19	1.0	68667	2	AC101811	AC101811 Mus muscu
146	20	1.0	108885	9	AC023135	AC023135 Homo sapi	c 219	19	1.0	73454	9	AL136988	AL136988 Human DNA

220	19	1.0	85097	3	AC026301	AC026301	Caenorhab	293	19	1.0	183775	2	AC093772	Homo sapi
221	19	1.0	86950	8	AL004218	AL004218	Arabidops	c 294	19	1.0	185272	2	AC090734	Homo sapi
222	19	1.0	89064	2	AL645826	AL645826	Danio rer	c 295	19	1.0	185528	2	AC020803	Mus muscu
223	19	1.0	93818	2	AL627423	Danio rer	296	19	1.0	185785	3	AC091221	Drosophil	
224	19	1.0	94236	2	AC020539	Drosophil	297	19	1.0	187276	2	AC104247	Homo sapi	
225	19	1.0	101614	2	AC008603	Homo sapi	c 298	19	1.0	188105	9	AC073655	Homo sapi	
226	19	1.0	110000	2	AC092450_3	Continuation (4 of	299	19	1.0	189036	2	AC019142	Homo sapi	
227	19	1.0	110000	2	AC026673_0	Continuation (4 of	300	19	1.0	190039	2	AC087639	Homo sapi	
228	19	1.0	110785	9	AC074183	Homo sapi	301	19	1.0	190197	2	AC026381	Mus muscu	
229	19	1.0	117157	8	AC034258	Oryza sat	302	19	1.0	191162	2	AC096629	Bos tauru	
230	19	1.0	118873	9	HS1184F4	Human DNA	c 303	19	1.0	191601	2	AC092143	Homo sapi	
231	19	1.0	123530	2	AL356369	Homo sapi	304	19	1.0	192021	2	AP003536	Oryza sat	
232	19	1.0	124280	9	HS03393J4	Human DNA	305	19	1.0	194350	10	AC079832	Mus muscu	
233	19	1.0	124758	2	AC008464	Homo sapi	c 306	19	1.0	197646	2	AC011041	Homo sapi	
234	19	1.0	125231	9	AC027296	Homo sapi	c 307	19	1.0	198259	9	AC016825	Homo sapi	
235	19	1.0	126126	2	AC093304	Homo sapi	c 308	19	1.0	198616	2	AC093139	Pan trogl	
236	19	1.0	131928	9	AC091738	Homo sapi	c 309	19	1.0	199008	2	AL645668	Mus muscu	
237	19	1.0	132887	9	AL160408	Human DNA	310	19	1.0	205050	1	AJ414155	Yersinia	
238	19	1.0	133350	9	AC022511	Homo sapi	311	19	1.0	206117	2	AC079845	Mus muscu	
239	19	1.0	133974	2	AP004359	Oryza sat	312	19	1.0	207897	2	AL590389	Mus muscu	
240	19	1.0	138373	2	AC011380	Homo sapi	313	19	1.0	212029	2	AC011318	Homo sapi	
241	19	1.0	141293	2	AP004337	Oryza sat	c 314	19	1.0	225943	2	AL645934	Mus muscu	
242	19	1.0	142348	9	AP003788	Homo sapi	c 315	19	1.0	237405	2	AC068142	Mus muscu	
243	19	1.0	148336	9	AC007110	Homo sapi	316	19	1.0	240968	2	AC021709	Mus muscu	
244	19	1.0	150263	9	AC073218	Homo sapi	317	19	1.0	247696	2	AC007493	Homo sapi	
245	19	1.0	151008	2	AP004261	Oryza sat	c 318	19	1.0	251039	2	AL592290	Homo sapi	
246	19	1.0	152185	2	AC012219	Homo sapi	319	19	1.0	252360	2	AC091272	Mus muscu	
247	19	1.0	152794	9	AL139376	Human DNA	320	19	1.0	253176	3	AE003718	Drosophil	
248	19	1.0	153087	2	AC017072	Homo sapi	321	19	1.0	280464	2	AC021061	Mus muscu	
249	19	1.0	153704	9	AC008567	Homo sapi	322	19	1.0	298216	2	AC006875	Caenorhab	
250	19	1.0	155086	2	AC109913	Bos tauru	323	19	1.0	331326	2	AC006802	Caenorhab	
251	19	1.0	155276	2	AC091992	Homo sapi	324	19	1.0	339650	1	AP003583	Nostoc sp	
252	19	1.0	155334	2	AC015822	Homo sapi	325	19	1.0	346296	8	AF090447	Zea mays	
253	19	1.0	155827	2	AC104172	Homo sapi	326	18	0.9	18	6	AB3831	Sequence 9	
254	19	1.0	156077	9	AL354997	Human DNA	327	18	0.9	18	6	AB3832	Sequence 10	
255	19	1.0	159185	2	AC074218	Mus muscu	c 328	18	0.9	21	6	AX110393	Sequence 11	
256	19	1.0	159499	2	AL645612	Homo sapi	329	18	0.9	24	6	AX110462	Sequence 12	
257	19	1.0	159664	2	AL592441	Homo sapi	c 330	18	0.9	70	10	MMU278764	Sequence 13	
258	19	1.0	160824	9	AC079228	Homo sapi	c 331	18	0.9	306	5	AY046292	Podarcis	
259	19	1.0	162036	2	AC015773	Homo sapi	332	18	0.9	352	5	AB061613	Gallus ga	
260	19	1.0	163469	9	AC080009	Homo sapi	333	18	0.9	361	5	AB003864	Gallus ga	
261	19	1.0	163803	9	AC093531	Homo sapi	334	18	0.9	381	11	G41116	G41116 Z6403 zebra	
262	19	1.0	164721	2	AC013657	Homo sapi	335	18	0.9	390	11	G55837	SHGC-101236	
263	19	1.0	164798	2	AC098686	Bos tauru	336	18	0.9	485	11	G65254	SHGC-101236	
264	19	1.0	165067	9	AL138927	Human DNA	337	18	0.9	512	6	AX261418	Sequence 14	
265	19	1.0	167075	2	AC092858	Bos tauru	338	18	0.9	534	11	G51059	SHGC-79698	
266	19	1.0	167642	2	AC027235	Homo sapi	c 339	18	0.9	568	8	AF307005	Leucoma	
267	19	1.0	169660	9	AY050668	Homo sapi	340	18	0.9	584	11	G37476	SHGC-57692	
268	19	1.0	169660	9	AC011479	Homo sapi	341	18	0.9	629	8	PE278482	P-exul 5.8S	
269	19	1.0	170440	2	AC048375	Homo sapi	c 342	18	0.9	700	6	AX183474	Sequence 15	
270	19	1.0	170944	2	AC094217	Rattus no	c 343	18	0.9	720	8	CNS01B2A	Botrytis	
271	19	1.0	171210	2	AC087543	Oryza sat	c 344	18	0.9	755	6	AX008796	Sequence 16	
272	19	1.0	171210	2	AC025905	Oryza sat	c 345	18	0.9	772	1	AF462019	Vibrio ch	
273	19	1.0	171617	2	AC016053	Homo sapi	c 346	18	0.9	855	8	ATH251851	Arabidops	
274	19	1.0	172339	2	AC092873	Pan trogl	c 347	18	0.9	880	8	D38020	Beta trigyn	
275	19	1.0	172983	9	AC002454	Human BAC	c 348	18	0.9	932	6	A98054	Sequence 17	
276	19	1.0	173807	2	AC107383	Homo sapi	c 349	18	0.9	950	3	AF162796	Sequence 18	
277	19	1.0	173997	2	AC012340	Homo sapi	c 350	18	0.9	979	8	TOBRUBPA	Drosophil	
278	19	1.0	174296	2	AC013805	Homo sapi	c 351	18	0.9	1077	5	AF206531	Tobacco rib	
279	19	1.0	174415	2	AC104666	Homo sapi	c 352	18	0.9	1079	8	AY045976	Arabidops	
280	19	1.0	175440	3	AC007804	Drosophil	353	18	0.9	1117	4	D86219	Oryctolagus	
281	19	1.0	175789	2	AP000425	Homo sapi	c 354	18	0.9	1171	3	CEPRP1P1	Arabidops	
282	19	1.0	176218	2	AC107424	Homo sapi	c 355	18	0.9	1192	3	AF151918	Caenorhab	
283	19	1.0	179007	2	AC012275	Homo sapi	c 356	18	0.9	1193	8	D38021	Pycnoporu	
284	19	1.0	179057	2	AC108677	Homo sapi	c 357	18	0.9	1205	3	AF151917	Beta webbii	
285	19	1.0	179132	10	AL358892	Mouse DNA	c 358	18	0.9	1208	3	AF151921	Pycnoporu	
286	19	1.0	179364	2	AC105447	Homo sapi	c 359	18	0.9	1209	3	AF151920	Prosthode	
287	19	1.0	180485	2	AC027145	Homo sapi	c 360	18	0.9	1257	9	AF131803	Prosthode	
288	19	1.0	180672	9	AC011451	Homo sapi	c 361	18	0.9	1423	8	D38017	Homo sapi	
289	19	1.0	181192	2	AC106085	Rattus no	362	18	0.9	1427	9	D38017	Beta trigyn	
290	19	1.0	181395	9	AC083800	Homo sapi	363	18	0.9	1429	5	AB006536	Homo sapi	
291	19	1.0	182080	2	AC079682	Mus muscu	c 364	18	0.9	1429	5	XIPNE6G	Xiphophorus	
292	19	1.0	182435	9	AC096591	Homo sapi	c 365	18	0.9	1429	6	A46744	Xiphophorus	

c 366	18	0.9	1643	8	D38018	D38018 Beta webbia	439	18	0.9	45005	9	AC004511	AC004511 Homo sapi
c 367	18	0.9	1697	8	CRSAHH1	Z26881 C.roseus SA	c 440	18	0.9	46509	2	AC013091	AC013091 Drosophill
c 368	18	0.9	1975	9	AK022414	AK022414 Homo sapi	c 441	18	0.9	49347	2	AC014413	AC014413 Drosophill
c 369	18	0.9	1980	8	SCU22449	U22449 Saccharomyc	c 442	18	0.9	49959	6	AX015907	AX015907 Sequence
c 370	18	0.9	2040	8	D38016	D38016 Sugar beet	c 443	18	0.9	50323	3	AE003118	AE003118 Drosophill
c 371	18	0.9	2048	1	AF274062	AF274062 Lactobaci	444	18	0.9	50974	9	AP000958	AP000958 Homo sapi
c 372	18	0.9	2068	9	F202550S01	F202550 Homo sapi	445	18	0.9	51211	2	AC107728	AC107728 Mus muscu
c 373	18	0.9	2073	8	BVRFPLPS	X87637 B.vulgaris	446	18	0.9	52149	2	AC101590	AC101590 Homo sapi
c 374	18	0.9	2074	8	BVRFPLPF	X87636 B.vulgaris	447	18	0.9	52520	9	AL591592	AL591592 Human DNA
c 375	18	0.9	2083	4	OCU35428	U35428 Oryctolagus	c 448	18	0.9	54039	9	AL606803	AL606803 Human DNA
c 376	18	0.9	2089	1	ECENO	X82400 E.coli eno	c 449	18	0.9	57906	2	AC108909	AC108909 Mus muscu
c 377	18	0.9	2097	9	ECNO1309	ECNO1309 Homo sapi	c 450	18	0.9	59289	2	AC108629	AC108629 Rattus no
c 378	18	0.9	2121	8	SCYKL074C	Z28074 S.cerevisia	c 451	18	0.9	59429	2	AC091176	AC091176 Homo sapi
c 379	18	0.9	2215	3	AY069334	AY069334 Drosophill	c 452	18	0.9	60019	3	AC005650	AC005650 Drosophill
c 380	18	0.9	2293	8	NPRBCS8B	X13711 Nicotiana p	c 453	18	0.9	60379	2	AC096619	AC096619 Rattus no
c 381	18	0.9	2293	8	TOBRBCS8B	M36685 B.plumbagin	c 454	18	0.9	60761	2	AC102704	AC102704 Mus muscu
c 382	18	0.9	2391	1	BACADESYN	M36890 Bacillus su	c 455	18	0.9	62916	8	AB046427	AB046427 Arabidops
c 383	18	0.9	2500	14	D83410	D83410 Barley mild	456	18	0.9	63853	2	AC087663	AC087663 Homo sapi
c 384	18	0.9	2687	10	AB037362S1	AB037362 Mus muscu	c 457	18	0.9	63966	9	AL451141	AL451141 Human DNA
c 385	18	0.9	2700	8	AF106950	AF106950 Filobasid	c 458	18	0.9	64252	2	AC073032	AC073032 Homo sapi
c 386	18	0.9	2958	10	RNITI1H4	Y11283 R.norvegicu	c 459	18	0.9	64965	9	AC006962	AC006962 Homo sapi
c 387	18	0.9	3275	8	D38019	D38019 Sugar beet	460	18	0.9	66542	9	AF134576	AF134576 Homo sapi
c 388	18	0.9	3464	14	BMNA1	D10947 Barley mild	461	18	0.9	66673	9	AL138847	AL138847 Human DNA
c 389	18	0.9	3485	6	E06133	E06133 cDNA sequen	462	18	0.9	67292	2	AC102708	AC102708 Mus muscu
c 390	18	0.9	4350	8	YSCRNASE	L06209 Saccharomyc	463	18	0.9	67623	2	AC027436	AC027436 Homo sapi
c 391	18	0.9	4371	1	AF390867	AF390867 Paracoccu	c 464	18	0.9	67623	2	AC027436	AC027436 Homo sapi
c 392	18	0.9	5142	8	MIATPA12	X15982 Ascobolus i	c 465	18	0.9	68217	9	AL590664	AL590664 Human DNA
c 393	18	0.9	5578	3	AF162798	AF162798 Drosophill	c 466	18	0.9	68792	2	AC107724	AC107724 Mus muscu
c 394	18	0.9	5941	14	AF036333	AF036333 Diatraea	c 467	18	0.9	70714	2	AC101499	AC101499 Mus muscu
c 395	18	0.9	6887	6	AX344607	AX344607 Sequence	c 468	18	0.9	71959	2	AC108436	AC108436 Mus muscu
c 396	18	0.9	7067	6	AX286953	AX286953 Sequence	c 469	18	0.9	73038	2	AC020884_3	Continuation (4 of
c 397	18	0.9	7187	2	AC014572	AC014572 Drosophill	c 470	18	0.9	73096	2	AC091063	AC091063 Homo sapi
c 398	18	0.9	7263	14	D83408	D83408 Barley mild	c 471	18	0.9	73390	9	AC012038	AC012038 Homo sapi
c 399	18	0.9	7647	8	AB027617	AB027617 Pyrus pyr	472	18	0.9	74508	5	AL591511	AL591511 Zebrafish
c 400	18	0.9	7925	3	AB042120	AB042120 Bombyx mo	c 473	18	0.9	74629	2	AC101440	AC101440 Mus muscu
c 401	18	0.9	7975	2	AC020446	AC020446 Drosophill	474	18	0.9	75031	9	AL390716	AL390716 Human DNA
c 402	18	0.9	8011	2	AC017787	AC017787 Drosophill	475	18	0.9	76388	2	AC100200	AC100200 Mus muscu
c 403	18	0.9	9347	9	HSIDAI56	Y15909 Homo sapien	476	18	0.9	79598	2	AC012281	AC012281 Homo sapi
c 404	18	0.9	9905	2	AC014733	AC014733 Drosophill	c 477	18	0.9	82912	8	ATT21J18	AL132963 Arabidops
c 405	18	0.9	10006	14	AF131870	AF131870 Simian im	c 478	18	0.9	83057	3	AL445703	AL445703 Human DNA
c 406	18	0.9	10050	1	AE004335	AE004335 Vibrio ch	c 479	18	0.9	83495	3	AC087076	AC087076 Caenorhab
c 407	18	0.9	10056	1	AE006480	AE006480 Streptoco	c 480	18	0.9	84196	8	ATT3A5	AL132979 Arabidops
c 408	18	0.9	10059	1	AE004369	AE004369 Vibrio ch	c 481	18	0.9	84392	2	AC013393	AC013393 Homo sapi
c 409	18	0.9	10330	1	AE006776	AE006776 Sulfolobu	c 482	18	0.9	84974	8	AC007843	AC007843 Arabidops
c 410	18	0.9	10385	1	AE000361	AE000361 Escherich	c 483	18	0.9	85186	2	AC021348	AC021348 Homo sapi
c 411	18	0.9	10707	1	AE002316	AE002316 Chlamydia	c 484	18	0.9	86075	8	NC93G11	AL513443 Neurospor
c 412	18	0.9	11142	1	AE000805	AE000805 Methanoba	c 485	18	0.9	86569	2	AC106139	AC106139 Rattus no
c 413	18	0.9	11311	1	AE005527	AE005527 Escherich	c 486	18	0.9	87397	9	AL356778	AL356778 Human DNA
c 414	18	0.9	11361	1	AE000950	AE000950 Archaeogl	c 487	18	0.9	87884	2	AC012997	AC012997 Drosophill
c 415	18	0.9	12111	1	AE006275	AE006275 Lactococc	c 488	18	0.9	89352	9	AC006030	AC006030 Homo sapi
c 416	18	0.9	12229	1	AE000913	AE000913 Methanoba	c 489	18	0.9	89732	2	AC091371	AC091371 Rattus no
c 417	18	0.9	13234	1	ECU29580	U29580 Escherichia	c 490	18	0.9	90141	9	AC035150	AC035150 Homo sapi
c 418	18	0.9	16071	1	AE000796	AE000796 Methanoba	c 491	18	0.9	90547	9	HS265J14	284477 Human DNA S
c 419	18	0.9	16238	1	AE007094	AE007094 Mycobacte	c 492	18	0.9	91817	2	AC109605	AC109605 Mus muscu
c 420	18	0.9	18400	3	CER13G10	Z35602 Caenorhabdi	c 493	18	0.9	92910	9	AL590368	AL590368 Human DNA
c 421	18	0.9	18689	1	AE001753	AE001753 Thermotog	c 494	18	0.9	93153	2	AF257499	AF257499 Homo sapi
c 422	18	0.9	19317	9	AL591862	AL591862 Human DNA	c 495	18	0.9	93217	2	AC023171	AC023171 Homo sapi
c 423	18	0.9	22700	2	AC020792	AC020792 Mus muscu	c 496	18	0.9	93893	9	AL669984	AL669984 Human DNA
c 424	18	0.9	23435	2	AC014167	AC014167 Drosophill	c 497	18	0.9	95091	9	AL353594	AL353594 Human DNA
c 425	18	0.9	23864	3	AE002853	AE002853 Drosophill	c 498	18	0.9	97011	9	HS997015	AL355073 Human DNA
c 426	18	0.9	23967	1	MTCY07A7	Z95556 Mycobacteri	c 499	18	0.9	97137	9	AC004745	AC004745 Homo sapi
c 427	18	0.9	29994	3	AE003032	AE003032 Drosophill	c 500	18	0.9	97700	2	HS280L10	AC031316 Human DNA
c 428	18	0.9	30352	3	U28412	U28412 Caenorhabdi	c 501	18	0.9	98334	2	AC095642	AC095642 Rattus no
c 429	18	0.9	32148	8	CEW04G3	Z46660 S.cerevisia	c 502	18	0.9	99228	9	AL133376	AL133376 Human DNA
c 430	18	0.9	32158	3	CEW04G3	Z68014 Caenorhabdi	c 503	18	0.9	100000	9	AP000066	AP000066 Homo sapi
c 431	18	0.9	32799	8	SPAC227	AL133156 S.pombe c	c 504	18	0.9	101188	2	AC010606	AC010606 Homo sapi
c 432	18	0.9	33330	8	AP001304	AP001304 Arabidops	c 505	18	0.9	101903	2	AP004334	AP004334 Oryza sat
c 433	18	0.9	35564	3	CEF23H12	Z74472 Caenorhabdi	c 506	18	0.9	103157	2	AC103199	AC103199 Rattus no
c 434	18	0.9	35935	9	AC092305	AC092305 Homo sapi	c 507	18	0.9	104475	9	AC011384	AC011384 Homo sapi
c 435	18	0.9	36127	3	CEF01D5	Z81493 Caenorhabdi	c 508	18	0.9	104995	9	AC020942	AC020942 Homo sapi
c 436	18	0.9	37577	8	AL590284	AL590284 Human DNA	c 509	18	0.9	104995	9	AL589653	AL589653 Human DNA
c 437	18	0.9	37949	8	SPBC725	AL034352 S.pombe c	c 510	18	0.9	105399	2	AC005427	AC005427 Drosophill
c 438	18	0.9	38976	3	AF063007	AF063007 Caenorhab	511	18	0.9	105931	2	AC005427	AC005427 Drosophill

c 512	18	0.9 107529	2	AC069085	Homo sapi	AC069085	18	c 585	18	0.9 149505	9	AC093767	Homo sapi	AC093767
c 513	18	0.9 108098	9	AC008411	Homo sapi	AC008411	18	c 586	18	0.9 149699	8	AP002486	Rattus sat	AP002486
c 514	18	0.9 108500	9	AP001624	Homo sapi	AP001624	18	c 587	18	0.9 149719	2	AC097312	Oryza sat	AC097312
c 515	18	0.9 109786	8	F5K24	Arabidops	AF128335	18	c 588	18	0.9 149745	2	AL016130	Drosophil	AL016130
c 516	18	0.9 110000	2	AC026673_0	Homo sapi	AC026673	18	c 589	18	0.9 149916	2	AL359540	Homo sapi	AL359540
c 517	18	0.9 111122	8	ATF11C1	Arabidops	AL132976	18	c 590	18	0.9 150424	9	AL391811	Human DNA	AL391811
c 518	18	0.9 111758	2	AC105893	Rattus no	AL105693	18	c 591	18	0.9 150606	2	AC027376	Homo sapi	AC027376
c 519	18	0.9 113194	2	AF301236	Homo sapi	AF301236	18	c 592	18	0.9 150997	9	HS11703	Human DNA	AL020995
c 520	18	0.9 113473	9	HS1040613	Human DNA	AL121887	18	c 593	18	0.9 151131	2	AP003916	Oryza sat	AP003916
c 521	18	0.9 115532	2	AC008424	Homo sapi	AC008424	18	c 594	18	0.9 151305	2	AC072029	Homo sapi	AC072029
c 522	18	0.9 115929	2	AC078901	Homo sapi	AC078901	18	c 595	18	0.9 151611	2	AP001131	Homo sapi	AP001131
c 523	18	0.9 116261	9	AL355357	Human DNA	AL355357	18	c 596	18	0.9 151679	2	AP001135	Homo sapi	AP001135
c 524	18	0.9 118068	2	AL356798	Human DNA	AL356798	18	c 597	18	0.9 151734	2	AC021225	Homo sapi	AC021225
c 525	18	0.9 118447	2	AC087397	Homo sapi	AC087397	18	c 598	18	0.9 151795	9	HS85661	Homo sapi	AC033361
c 526	18	0.9 118507	8	ATF7J8	Arabidops	AL137189	18	c 599	18	0.9 151900	9	AC096644	Human DNA	AC096644
c 527	18	0.9 118911	2	AC099063	Homo sapi	AC099063	18	c 600	18	0.9 151938	9	HS323M4	Human DNA	AL033378
c 528	18	0.9 119063	9	AC027311	Homo sapi	AC027311	18	c 601	18	0.9 151959	2	AC021689	Homo sapi	AC021689
c 529	18	0.9 119430	8	ATF9D16	Arabidops	AL035394	18	c 602	18	0.9 151987	2	AP004339	Oryza sat	AP004339
c 530	18	0.9 121496	2	AC090244	Homo sapi	AC090244	18	c 603	18	0.9 152269	2	AL591024	Homo sapi	AL591024
c 531	18	0.9 122250	9	AC005273	Homo sapi	AC005273	18	c 604	18	0.9 152336	9	AP004714	Homo sapi	AP004714
c 532	18	0.9 123010	9	AC087321	Homo sapi	AC087321	18	c 605	18	0.9 152354	2	AC095280	Rattus no	AC095280
c 533	18	0.9 123519	30	AL161639	Human DNA	AL161639	18	c 606	18	0.9 152702	2	AC106396	Homo sapi	AC106396
c 534	18	0.9 123642	2	AC094853	Rattus no	AC094853	18	c 607	18	0.9 152805	9	AC078923	Homo sapi	AC078923
c 535	18	0.9 123805	9	AC002464	Human BAC	AC002464	18	c 608	18	0.9 152837	2	AC105190	Homo sapi	AC105190
c 536	18	0.9 123837	2	AC013661	Homo sapi	AC013661	18	c 609	18	0.9 152951	9	AC026371	Homo sapi	AC026371
c 537	18	0.9 123953	2	AP003747	Oryza sat	AP003747	18	c 610	18	0.9 152996	9	AC013719	Homo sapi	AC013719
c 538	18	0.9 124368	2	AC055703	Mus muscu	AC055703	18	c 611	18	0.9 153528	2	AC024031	Homo sapi	AC024031
c 539	18	0.9 125045	2	AC094470	Rattus no	AC094470	18	c 612	18	0.9 154011	2	AC018571	Homo sapi	AC018571
c 540	18	0.9 125685	9	AC073310	Homo sapi	AC073310	18	c 613	18	0.9 154189	2	AC073180	Homo sapi	AC073180
c 541	18	0.9 126589	9	AL512444	Human DNA	AL512444	18	c 614	18	0.9 154588	9	AL355871	Human DNA	AL355871
c 542	18	0.9 126619	9	AC010265	Homo sapi	AC010265	18	c 615	18	0.9 154772	9	AL138764	Human DNA	AL138764
c 543	18	0.9 126910	9	AC005856	Homo sapi	AC005856	18	c 616	18	0.9 154880	2	AC084695	Mus muscu	AC084695
c 544	18	0.9 127192	9	HS106502	Human DNA	AL035562	18	c 617	18	0.9 155191	2	AC087888	Homo sapi	AC087888
c 545	18	0.9 127577	2	AC104459	Homo sapi	AC104459	18	c 618	18	0.9 155304	2	AC024713	Homo sapi	AC024713
c 546	18	0.9 127968	8	AC022492	Genomic s	AC022492	18	c 619	18	0.9 155531	2	AL592213	Homo sapi	AL592213
c 547	18	0.9 130626	9	AP001625	Homo sapi	AP001625	18	c 620	18	0.9 155577	2	AC011565	Homo sapi	AC011565
c 548	18	0.9 130737	9	AC092360	Homo sapi	AC092360	18	c 621	18	0.9 155789	2	AL669860	Mus muscu	AL669860
c 549	18	0.9 131487	2	AP004117	Oryza sat	AP004117	18	c 622	18	0.9 155842	2	AC025650	Homo sapi	AC025650
c 550	18	0.9 131692	8	AC006423	Arabidops	AC006423	18	c 623	18	0.9 156503	9	AC011511	Homo sapi	AC011511
c 551	18	0.9 135438	9	AC026444	Homo sapi	AC026444	18	c 624	18	0.9 156741	2	AL355867	Homo sapi	AL355867
c 552	18	0.9 135752	9	AC073465	Homo sapi	AC073465	18	c 625	18	0.9 156836	2	AC068059	Homo sapi	AC068059
c 553	18	0.9 136095	2	AF161326	Homo sapi	AF161326	18	c 626	18	0.9 156870	2	AC091916	Homo sapi	AC091916
c 554	18	0.9 136328	9	AC004126	Human Chr	AC004126	18	c 627	18	0.9 157711	9	AC055738	Homo sapi	AC055738
c 555	18	0.9 136915	2	AC091028	Homo sapi	AC091028	18	c 628	18	0.9 158033	9	AC018926	Homo sapi	AC018926
c 556	18	0.9 136917	2	AC087220	Oryza sat	AC087220	18	c 629	18	0.9 158392	2	AC022048	Homo sapi	AC022048
c 557	18	0.9 137008	10	AC091618	Rattus no	AC091618	18	c 630	18	0.9 158408	9	AL391821	Human DNA	AL391821
c 558	18	0.9 137897	9	AL161913	Human DNA	AL161913	18	c 631	18	0.9 158723	8	AP002869	Oryza sat	AP002869
c 559	18	0.9 138053	9	AC091146	Homo sapi	AC091146	18	c 632	18	0.9 158952	9	AC096745	Homo sapi	AC096745
c 560	18	0.9 138473	10	AC003019	Mus muscu	AC003019	18	c 633	18	0.9 158991	2	AC087673	Homo sapi	AC087673
c 561	18	0.9 138752	9	AL161627	Human DNA	AL161627	18	c 634	18	0.9 158995	2	AL590679	Homo sapi	AL590679
c 562	18	0.9 138781	2	AP002764	Homo sapi	AP002764	18	c 635	18	0.9 159003	2	AC025255	Homo sapi	AC025255
c 563	18	0.9 139904	9	HS97D16	Human DNA	AL009179	18	c 636	18	0.9 159128	2	AC040958	Homo sapi	AC040958
c 564	18	0.9 140229	8	AP003143	Oryza sat	AP003143	18	c 637	18	0.9 159163	9	CNS05TCY	Human chr	AL355887
c 565	18	0.9 140529	2	AC011486	Homo sapi	AC011486	18	c 638	18	0.9 159387	2	AC084858	Homo sapi	AC084858
c 566	18	0.9 140554	10	AC000399	Genomic s	AC000399	18	c 639	18	0.9 159507	9	AC004985	Homo sapi	AC004985
c 567	18	0.9 140921	2	AC104800	Homo sapi	AC104800	18	c 640	18	0.9 159620	2	AP002442	Homo sapi	AP002442
c 568	18	0.9 141152	2	AP004343	Oryza sat	AP004343	18	c 641	18	0.9 159930	2	AL138757	Homo sapi	AL138757
c 569	18	0.9 141456	9	AL355300	Human DNA	AL355300	18	c 642	18	0.9 160038	9	AL138755	Human DNA	AL138755
c 570	18	0.9 142040	2	AC103268	Rattus no	AC103268	18	c 643	18	0.9 160048	10	AC002109	Genomic s	AC002109
c 571	18	0.9 142239	2	AC027635	Homo sapi	AC027635	18	c 644	18	0.9 160209	9	AL138815	Human DNA	AL138815
c 572	18	0.9 143029	2	AC016319	Homo sapi	AC016319	18	c 645	18	0.9 160210	9	AL392163	Human DNA	AL392163
c 573	18	0.9 143078	9	AL445463	Human DNA	AL445463	18	c 646	18	0.9 160666	2	AC105755	Homo sapi	AC105755
c 574	18	0.9 143302	2	AP004264	Oryza sat	AP004264	18	c 647	18	0.9 160754	9	AC006213	Homo sapi	AC006213
c 575	18	0.9 143655	2	AC010808	Homo sapi	AC010808	18	c 648	18	0.9 160855	2	AC011579	Homo sapi	AC011579
c 576	18	0.9 144723	10	AC005818	Mus muscu	AC005818	18	c 649	18	0.9 160972	3	AC007723	Drosophil	AC007723
c 577	18	0.9 145089	9	AC022418	Homo sapi	AC022418	18	c 650	18	0.9 161042	2	AC098497	Rattus no	AC098497
c 578	18	0.9 145749	2	AC091359	Rattus no	AC091359	18	c 651	18	0.9 161103	9	AL392163	Human DNA	AL392163
c 579	18	0.9 146261	2	AC097112	Oryza sat	AC097112	18	c 652	18	0.9 161245	2	AP001986	Homo sapi	AP001986
c 580	18	0.9 146889	9	AL358937	Human DNA	AL358937	18	c 653	18	0.9 161313	2	AC095508	Homo sapi	AC095508
c 581	18	0.9 147109	9	AL591043	Human DNA	AL591043	18	c 654	18	0.9 161549	9	AC073174	Rattus no	AC073174
c 582	18	0.9 147404	2	AC025759	Homo sapi	AC025759	18	c 655	18	0.9 161941	2	AL645968	Mus muscu	AL645968
c 583	18	0.9 148119	2	AC010600	Homo sapi	AC010600	18	c 656	18	0.9 161970	2	AC027706	Homo sapi	AC027706
c 584	18	0.9 148335	9	HS809F4	Human DNA	AL022400	18	c 657	18	0.9 162112	2	AC011572	Homo sapi	AC011572

c 658	18	0.9 162217	9	AC068043	AC068043 Homo sapi	c 731	18	0.9 176805	2	AC107073	AC107073 Homo sapi
c 659	18	0.9 162282	9	AC094095	AC094095 Homo sapi	c 732	18	0.9 176835	2	AC099470	AC099470 Rattus no
c 660	18	0.9 162489	9	AP001572	AP001572 Homo sapi	c 733	18	0.9 176967	2	AC022742	AC022742 Homo sapi
c 661	18	0.9 162587	9	AC092805	AC092805 Homo sapi	c 734	18	0.9 177073	2	AP000934	AP000934 Homo sapi
c 662	18	0.9 162616	9	AC099539	AC099539 Homo sapi	c 735	18	0.9 177390	2	AC096113	AC096113 Rattus no
c 663	18	0.9 162786	2	AC010783	AC010783 Homo sapi	c 736	18	0.9 178171	2	AC026961	AC026961 Homo sapi
c 664	18	0.9 162919	2	AC092088	AC092088 Canis fam	c 737	18	0.9 178282	2	AC105349	AC105349 Homo sapi
c 665	18	0.9 163110	2	AC068705	AC068705 Homo sapi	c 738	18	0.9 178405	9	AC034216	AC034216 Homo sapi
c 666	18	0.9 163426	2	AC098176	AC098176 Rattus no	c 739	18	0.9 178515	2	AC009621	AC009621 Homo sapi
c 667	18	0.9 163556	9	AL592220	AL592220 Human DNA	c 740	18	0.9 178581	2	HS512L9	AL445624 Homo sapi
c 668	18	0.9 163612	2	AC103253	AC103253 Rattus no	c 741	18	0.9 178791	3	AC008218	AC008218 Drosophill
c 669	18	0.9 163706	2	AL354927	AL354927 Homo sapi	c 742	18	0.9 178806	2	AC013557	AC013557 Homo sapi
c 670	18	0.9 163796	9	AC093725	AC093725 Homo sapi	c 743	18	0.9 178821	2	AC079640	AC079640 Mus muscu
c 671	18	0.9 163997	9	AC022731	AC022731 Homo sapi	c 744	18	0.9 178902	2	AL513525	AL513525 Homo sapi
c 672	18	0.9 164383	2	AC068023	AC068023 Homo sapi	c 745	18	0.9 178933	9	AL591479	AL591479 Human DNA
c 673	18	0.9 164950	2	AC022262	AC022262 Homo sapi	c 746	18	0.9 179199	2	AC106682	AC106682 Rattus no
c 674	18	0.9 165415	2	AC016663	AC016663 Homo sapi	c 747	18	0.9 179439	2	AC018874	AC018874 Homo sapi
c 675	18	0.9 166354	2	AC078984	AC078984 Homo sapi	c 748	18	0.9 179461	2	AC019177	AC019177 Homo sapi
c 676	18	0.9 166357	9	AC026307	AC026307 Homo sapi	c 749	18	0.9 179651	9	AC005553	AC005553 Homo sapi
c 677	18	0.9 166516	9	AC012068	AC012068 Homo sapi	c 750	18	0.9 180134	2	AL672022	AL672022 Homo sapi
c 678	18	0.9 166606	2	AC020685	AC020685 Homo sapi	c 751	18	0.9 180136	1	BAC180K	D26185 B. subtilis
c 679	18	0.9 166608	2	AC024727	AC024727 Homo sapi	c 752	18	0.9 180141	2	AC025824	AC025824 Homo sapi
c 680	18	0.9 166772	2	AC098691	AC098691 Homo sapi	c 753	18	0.9 180336	2	AC021731	AC021731 Homo sapi
c 681	18	0.9 166961	2	AC067990	AC067990 Homo sapi	c 754	18	0.9 180551	9	HUAC004158	AC004158 Homo sapi
c 682	18	0.9 166962	2	AC098180	AC098180 Rattus no	c 755	18	0.9 180990	2	AC097421	AC097421 Rattus no
c 683	18	0.9 167185	2	AC080108	AC080108 Homo sapi	c 756	18	0.9 181016	2	AC073129	AC073129 Homo sapi
c 684	18	0.9 167192	2	AC097301	AC097301 Rattus no	c 757	18	0.9 181567	9	AL354811	AL354811 Human DNA
c 685	18	0.9 167296	2	AP002507	AP002507 Homo sapi	c 758	18	0.9 181835	2	AC026124	AC026124 Homo sapi
c 686	18	0.9 167497	9	AL133479	AL133479 Human DNA	c 759	18	0.9 181994	2	AC090822	AC090822 Homo sapi
c 687	18	0.9 167650	2	HS519C2	AL449210 Homo sapi	c 760	18	0.9 182150	9	HSG278N14	AL109654 Human DNA
c 688	18	0.9 167782	2	AC010943	AC010943 Homo sapi	c 761	18	0.9 182340	2	AC024417	AC024417 Homo sapi
c 689	18	0.9 167965	2	AC095250	AC095250 Rattus no	c 762	18	0.9 182427	9	AL162330	AL162330 Human DNA
c 690	18	0.9 168265	9	AC013716	AC013716 Homo sapi	c 763	18	0.9 182770	9	AC104082	AC104082 Homo sapi
c 691	18	0.9 168837	9	AC090108	AC090108 Homo sapi	c 764	18	0.9 182839	2	AC022619	AC022619 Homo sapi
c 692	18	0.9 168991	9	AC073958	AC073958 Homo sapi	c 765	18	0.9 183060	9	AC018641	AC018641 Human Chr
c 693	18	0.9 169027	2	AC099282	AC099282 Rattus no	c 766	18	0.9 183181	8	ATCHRIV19	AL161507 Arabidops
c 694	18	0.9 169460	2	AC079068	AC079068 Homo sapi	c 767	18	0.9 183430	9	AL391827	AL391827 Human DNA
c 695	18	0.9 169476	9	AC016207	AC016207 Homo sapi	c 768	18	0.9 183515	9	AC016908	AC016908 Homo sapi
c 696	18	0.9 169566	2	AC073579	AC073579 Homo sapi	c 769	18	0.9 183903	9	AP003403	AP003403 Homo sapi
c 697	18	0.9 169936	9	CNS05TDD	AL356032 Human chr	c 770	18	0.9 183960	9	AC013445	AC013445 Homo sapi
c 698	18	0.9 170022	2	AC021559	AC021559 Homo sapi	c 771	18	0.9 184271	3	AC009256	AC009256 Drosophill
c 699	18	0.9 170090	2	AC096640	AC096640 Homo sapi	c 772	18	0.9 184873	2	AC026325	AC026325 Homo sapi
c 700	18	0.9 170479	2	AC016845	AC016845 Homo sapi	c 773	18	0.9 184925	2	AC027608	AC027608 Homo sapi
c 701	18	0.9 171216	2	AC025338	AC025338 Homo sapi	c 774	18	0.9 184938	2	AP002476	AP002476 Homo sapi
c 702	18	0.9 171621	2	AC105937	AC105937 Homo sapi	c 775	18	0.9 184999	2	AL611934	AL611934 Mus muscu
c 703	18	0.9 171962	2	AC013537	AC013537 Homo sapi	c 776	18	0.9 185134	9	AC007878	AC007878 Homo sapi
c 704	18	0.9 171985	9	AL445675	AL445675 Human DNA	c 777	18	0.9 185281	2	AC011214	AC011214 Homo sapi
c 705	18	0.9 172252	2	AC093950	AC093950 Homo sapi	c 778	18	0.9 185380	2	AC016863	AC016863 Homo sapi
c 706	18	0.9 172378	2	AC021696	AC021696 Homo sapi	c 779	18	0.9 185448	2	AC092549	AC092549 Homo sapi
c 707	18	0.9 172549	2	AC027519	AC027519 Homo sapi	c 780	18	0.9 185556	9	AC019133	AC019133 Homo sapi
c 708	18	0.9 172550	2	AC072035	AC072035 Homo sapi	c 781	18	0.9 185654	2	AC069564	AC069564 Mus muscu
c 709	18	0.9 172786	2	AC105379	AC105379 Trypanoso	c 782	18	0.9 185664	10	AC006507	AC006507 Mus muscu
c 710	18	0.9 172798	9	AC026585	AC026585 Homo sapi	c 783	18	0.9 186253	9	AC020892	AC020892 Homo sapi
c 711	18	0.9 172828	9	AC009948	AC009948 Homo sapi	c 784	18	0.9 186519	2	AC073476	AC073476 Homo sapi
c 712	18	0.9 173119	9	AL359986	AL359986 Human DNA	c 785	18	0.9 186526	2	AC108067	AC108067 Homo sapi
c 713	18	0.9 173613	3	AC007475	AC007475 Drosophill	c 786	18	0.9 186540	9	AC012478	AC012478 Homo sapi
c 714	18	0.9 173792	30	AC022216	AC022216 Homo sapi	c 787	18	0.9 186919	2	AC091544	AC091544 Homo sapi
c 715	18	0.9 173947	9	AF274856	AF274856 Homo sapi	c 788	18	0.9 187038	2	AC021179	AC021179 Homo sapi
c 716	18	0.9 174487	2	AL662924	AL662924 Homo sapi	c 789	18	0.9 187517	9	HSBA192p3	AL117340 Human DNA
c 717	18	0.9 174579	2	AC016666	AC016666 Homo sapi	c 790	18	0.9 187621	2	AC079060	AC079060 Homo sapi
c 718	18	0.9 174624	2	AL160396	AL160396 Homo sapi	c 791	18	0.9 187914	9	AC104126	AC104126 Homo sapi
c 719	18	0.9 174716	2	AC094819	AC094819 Rattus no	c 792	18	0.9 189058	9	AC015724	AC015724 Homo sapi
c 720	18	0.9 174907	2	AC025455	AC025455 Homo sapi	c 793	18	0.9 189095	9	AL512430	AL512430 Human DNA
c 721	18	0.9 175271	9	AL354718	AL354718 Human DNA	c 794	18	0.9 189229	2	AC090138	AC090138 Homo sapi
c 722	18	0.9 175451	3	AC099035	AC099035 Drosophill	c 795	18	0.9 189290	2	AC106165	AC106165 Rattus no
c 723	18	0.9 175488	2	AC013556	AC013556 Homo sapi	c 796	18	0.9 189667	2	AC079553	AC079553 Mus muscu
c 724	18	0.9 175494	9	AL590667	AL590667 Human DNA	c 797	18	0.9 190174	30	AC022739	AC022739 Homo sapi
c 725	18	0.9 175563	2	AC096341	AC096341 Rattus no	c 798	18	0.9 190221	2	AC025741	AC025741 Homo sapi
c 726	18	0.9 175752	2	AC090117	AC090117 Homo sapi	c 799	18	0.9 190418	2	AP000777	AP000777 Homo sapi
c 727	18	0.9 175782	2	AL627074	AL627074 Mus muscu	c 800	18	0.9 190766	2	AC087840	AC087840 Mus muscu
c 728	18	0.9 175967	9	AL591438	AL591438 Human DNA	c 801	18	0.9 190793	2	AC034298	AC034298 Homo sapi
c 729	18	0.9 176097	9	AL591438	AL591438 Human DNA	c 802	18	0.9 191115	2	AC069029	AC069029 Homo sapi
c 730	18	0.9 176678	30	AC036213	AC036213 Homo sapi	c 803	18	0.9 191438	3	AC008351	AC008351 Drosophill

804	18	0.9	191583	9	AC015651	AC015651 Homo sapi	877	18	0.9	240825	6	AX087869	AX087869 Sequence
805	18	0.9	191704	2	AC053529	AC053529 Homo sapi	c 878	18	0.9	259109	2	AL591851	AL591851 Homo sapi
806	18	0.9	191853	2	AC084151	AC084151 Homo sapi	c 879	18	0.9	259894	2	HUAC002302	HUAC002302 Homo sapi
c 807	18	0.9	192440	9	CNS01DXH	AL139316 Human chr	880	18	0.9	262731	3	AE003823	AE003823 Drosophila
c 808	18	0.9	192730	2	AC026676	AC026676 Homo sapi	881	18	0.9	266558	1	AP002563	AP002563 Escherich
809	18	0.9	192763	3	AC007474	AC007474 Drosophila	c 882	18	0.9	295054	2	AC068887	AC068887 Homo sapi
810	18	0.9	193096	2	AC104395	AC104395 Homo sapi	883	18	0.9	299751	3	AE003459	AE003459 Drosophila
c 811	18	0.9	193594	2	AL662903	AL662903 Homo sapi	884	18	0.9	309207	3	AE003501	AE003501 Drosophila
812	18	0.9	193622	2	AC009659	AC009659 Homo sapi	885	18	0.9	305518	2	AP006879	AP006879 Caenorhab
813	18	0.9	193925	2	AC093560	AC093560 Homo sapi	886	18	0.9	340000	9	AP001675	AP001675 Homo sapi
814	18	0.9	194297	2	AL607129	AL607129 Mus muscu	c 887	18	0.9	340000	9	AP001746	AP001746 Homo sapi
815	18	0.9	194520	9	AL356915	AL356915 Human DNA	c 888	18	0.9	348250	1	AP003592	AP003592 Nostoc sp
816	18	0.9	194601	2	AC079949	AC079949 Homo sapi	889	18	0.9	27	6	A98067	A98067 Sequence 14
c 817	18	0.9	194634	2	AC020286	AC020286 Drosophila	890	18	0.9	27	6	A98074	A98074 Sequence 21
818	18	0.9	194823	2	AC087369	AC087369 Homo sapi	c 891	18	0.9	27	6	AR120352	AR120352 Sequence
c 819	18	0.9	194946	2	AC027362	AC027362 Homo sapi	c 892	18	0.9	35	6	E39903	E39903 High-affini
820	18	0.9	195111	2	AC096981	AC096981 Rattus no	c 893	18	0.9	36	6	E39902	E39902 High-affini
c 821	18	0.9	195472	2	AC018398	AC018398 Homo sapi	c 894	18	0.9	145	11	HUMSWX1127	L34349 Human chrom
822	18	0.9	195616	2	AC090344	AC090344 Homo sapi	895	18	0.9	207	11	KLAJ9770	AJ229770 Kluyverom
823	18	0.9	195743	2	AC016774	AC016774 Homo sapi	896	18	0.9	302	3	S78746	S78746 h-hairy pai
824	18	0.9	196349	2	AC022032	AC022032 Homo sapi	c 897	18	0.9	306	5	AY046281	AY046281 Podarcis
c 825	18	0.9	196349	2	AC022032	AC022032 Homo sapi	c 898	18	0.9	306	5	AY046283	AY046283 Podarcis
826	18	0.9	196677	8	ATC8RIV60	AC073479 Homo sapi	c 899	18	0.9	306	5	AY046284	AY046284 Podarcis
827	18	0.9	197156	2	AC090345	AL161560 Arabidops	c 900	18	0.9	306	5	AY046286	AY046286 Podarcis
828	18	0.9	197322	9	AL592144	AC090345 Homo sapi	c 901	18	0.9	321	10	MMIGD4	V00789 Mouse gene
829	18	0.9	197348	3	AC008028	AL592144 Human DNA	902	18	0.9	324	11	G71383	G71383 716234231FM
c 830	18	0.9	198084	9	AF225898	AC008028 Drosophila	903	18	0.9	332	8	ABE291782	AJ291782 Atropa be
c 831	18	0.9	198241	2	AP003721	AF225898 Homo sapi	904	18	0.9	358	5	AB061589	AB061589 Gallus ga
c 832	18	0.9	198523	2	AC105807	AP003721 Homo sapi	905	18	0.9	362	5	D63959	D63959 Gallus gall
c 833	18	0.9	199000	8	AF000396S1	AC105807 Rattus no	906	18	0.9	364	5	AB003845	AB003845 Gallus ga
834	18	0.9	199199	8	ATC8RIV59	AP000396 Beta vulg	907	18	0.9	367	5	AF339339	AF339339 Anas plat
c 835	18	0.9	199515	2	AC027705	AL161559 Arabidops	c 908	18	0.9	368	11	BTM428564	AJ2428564 Bos tauru
c 836	18	0.9	199531	9	AC094104	AC027705 Homo sapi	c 909	18	0.9	376	11	G50078	G50078 SHGC-82970
837	18	0.9	199636	2	AL645855	AC094104 Homo sapi	910	18	0.9	377	5	CHKIGLA14	M15144 Chicken Ig
c 838	18	0.9	199928	2	AC092500	AL645855 Mus muscu	c 911	18	0.9	393	9	HSCD30L5	AF006385 Homo sapi
c 839	18	0.9	200500	2	AC093538	AC092500 Homo sapi	c 912	18	0.9	396	4	AF028145	AF028145 Dendrocyon
c 840	18	0.9	200582	2	AC008746	AC093538 Homo sapi	c 913	18	0.9	405	3	AF191165	AF191165 Drosomyr
c 841	18	0.9	200799	2	AL445648	AC008746 Homo sapi	c 914	18	0.9	405	9	AF274755S55	AF274755 Homo sapi
c 842	18	0.9	200909	2	AC103915	AL445648 Homo sapi	c 915	18	0.9	419	1	SPPBFLAB	Z21809 S.pneumonia
843	18	0.9	202929	3	AC007892	AC103915 Homo sapi	916	18	0.9	425	11	G53163	G53163 SHGC-84696
c 844	18	0.9	205327	2	AC107396	AC007892 Drosophila	917	18	0.9	451	3	U91383	U91383 Rivellia qu
c 845	18	0.9	205860	9	CNS01RGN	AC107396 Homo sapi	c 918	18	0.9	468	11	G64794	G64794 B141616 GSS
c 846	18	0.9	206258	2	AC073916	AL158801 Human chr	c 919	18	0.9	494	1	L51APD	M80347 Listeria im
c 847	18	0.9	207945	9	CNS01DSS	AC073916 Homo sapi	c 920	18	0.9	495	10	MMU296878	AJ296878 Mus muscu
848	18	0.9	208729	2	AC012052	AL121841 Human chr	921	18	0.9	498	8	PDY15429	Y15429 Populus del
c 849	18	0.9	210350	2	AC079536	AC012052 Homo sapi	c 922	18	0.9	511	6	AX332740	AX332740 Sequence
c 850	18	0.9	210400	2	AC108141	AC079536 Mus muscu	c 923	18	0.9	511	6	AX332912	AX332912 Sequence
c 851	18	0.9	211127	2	AC067856	AC108141 Homo sapi	c 924	18	0.9	537	5	LP1249753	AJ249753 Lygodacty
c 852	18	0.9	211628	2	AC099425	AC067856 Homo sapi	c 925	18	0.9	560	1	BCE243161	AJ243161 Bacillus
853	18	0.9	212360	2	AL606908	AC099425 Rattus no	c 926	18	0.9	560	1	BCE243162	AJ243162 Bacillus
854	18	0.9	213447	2	AL354875	AL606908 Mus muscu	c 927	18	0.9	560	1	BCE243163	AJ243163 Bacillus
855	18	0.9	214192	2	AL646044	AL354875 Homo sapi	c 928	18	0.9	560	1	BTH243147	AJ243147 Bacillus
856	18	0.9	215177	2	AC011667	AL646044 Mus muscu	c 929	18	0.9	560	1	BTH243149	AJ243149 Bacillus
c 857	18	0.9	215534	1	BSUH0021	AC011667 Homo sapi	930	18	0.9	560	6	AX111540	AX111540 Sequence
c 858	18	0.9	215717	2	AC092196	Z99124 Bacillus su	931	18	0.9	560	6	AX111542	AX111542 Sequence
c 859	18	0.9	218233	2	AC099669	AC092196 Canis fam	c 932	18	0.9	564	14	HIM404104	AJ404104 Human imm
860	18	0.9	218286	2	AC099727	AC099669 Homo sapi	933	18	0.9	601	5	AF339325	AF339325 Pelaeuice
861	18	0.9	218807	2	AC009175	AC099727 Mus muscu	934	18	0.9	602	5	AF339324	AF339324 Balanice
862	18	0.9	218941	2	AC092950	AC009175 Homo sapi	935	18	0.9	603	1	AB036433	AB036433 Unculture
863	18	0.9	219180	2	AC092251	AC092950 Homo sapi	c 936	18	0.9	616	8	AF231284	AF231284 Leucoloma
c 864	18	0.9	220260	2	AC103535	AC092251 Mus muscu	c 937	18	0.9	616	8	AF231285	AF231285 Leucoloma
c 865	18	0.9	220699	2	AC009206	AC103535 Rattus no	c 938	18	0.9	616	8	AF231286	AF231286 Leucoloma
c 866	18	0.9	221276	2	AC098721	AC009206 Drosophila	c 939	18	0.9	646	8	AF382710	AF382710 Thibaudia
867	18	0.9	222235	3	AC074148	AC098721 Mus muscu	940	18	0.9	648	5	AF168056	AF168056 Opisthoco
868	18	0.9	224100	3	AE003770	AC074148 Mus muscu	941	18	0.9	657	10	MUSIGCD16	J00449 Mouse germi
c 869	18	0.9	226502	2	AL645802	AE003770 Drosophila	942	18	0.9	668	10	MUSIGKAG3	L00031 Mouse Ig ka
870	18	0.9	226916	9	AC005066	AL645802 Mus muscu	943	18	0.9	706	6	AX154845	AX154845 Sequence
c 871	18	0.9	228283	10	AF312994	AC005066 Homo sapi	944	18	0.9	718	3	AF146717	AF146717 Dorymyrme
c 872	18	0.9	229122	3	AE003700	AF312994 Mus muscu	945	18	0.9	722	5	GDSCABL1	Z46722 G.domesticu
c 873	18	0.9	229380	2	AC079636	AE003700 Drosophila	946	18	0.9	724	33	AC061416	AC061416 Giardia i
874	18	0.9	229518	2	AC069058	AC079636 Mus muscu	c 947	18	0.9	728	8	TOHMT206	D84426 Tomato mito
c 875	18	0.9	229518	2	AC069058	AC069058 Homo sapi	948	18	0.9	730	8	PHZ78451	Z78451 P.hookerae
876	18	0.9	239171	3	AE003765	AC069058 Homo sapi	c 949	18	0.9	752	11	G64645	G64645 B141A22/T7

c 950	17	0.9	783	6	A01796	A01796 Synthetic g
c 951	17	0.9	788	9	HSN801164	AL117628 Homo sapi
c 952	17	0.9	814	33	AC062780	Ac062780 Giardlia i
c 953	17	0.9	820	33	AC032294	Ac032294 Giardlia i
c 954	17	0.9	834	6	AX106224	AX106224 Sequence
c 955	17	0.9	834	6	AX140515	AX140515 Sequence
c 956	17	0.9	834	6	AX200375	AX200375 Sequence
c 957	17	0.9	834	6	AX267031	AX267031 Sequence
c 958	17	0.9	834	33	AC053681	Ac053681 Giardlia i
c 959	17	0.9	881	33	AC034903	Ac034903 Giardlia i
c 960	17	0.9	888	33	AC064268	Ac064268 Giardlia i
c 961	17	0.9	913	33	AC086182	Ac086182 Giardlia i
c 962	17	0.9	921	33	AC081013	Ac081013 Giardlia i
c 963	17	0.9	925	33	AC042131	Ac042131 Giardlia i
c 964	17	0.9	931	33	AC082542	Ac082542 Giardlia i
c 965	17	0.9	965	33	AC033632	Ac033632 Giardlia i
c 966	17	0.9	968	33	AC067501	Ac067501 Giardlia i
c 967	17	0.9	971	6	AR175273	AR175273 Sequence
c 968	17	0.9	972	6	AX349609	AX349609 Sequence
c 969	17	0.9	972	8	ANMP	X52565 A. nidulans
c 970	17	0.9	988	33	AC066160	Ac066160 Giardlia i
c 971	17	0.9	999	33	AC065913	Ac065913 Giardlia i
c 972	17	0.9	1000	8	TROPRTG	X55879 T. reesei u
c 973	17	0.9	1001	33	AC054213	Ac054213 Giardlia i
c 974	17	0.9	1003	5	AF125514	AF125514 Oncorhync
c 975	17	0.9	1003	5	AF125515	AF125515 Oncorhync
c 976	17	0.9	1005	14	SCL290434	AJ290434 Subterranean
c 977	17	0.9	1009	8	U01020	U01020 Medicago sa
c 978	17	0.9	1012	33	AC086876	Ac086876 Giardlia i
c 979	17	0.9	1031	33	AC043181	Ac043181 Giardlia i
c 980	17	0.9	1032	8	MIORF206	X74164 O. berterian
c 981	17	0.9	1036	5	AF044132	AF044132 Oncorhync
c 982	17	0.9	1036	5	AF044154	AF044154 Oncorhync
c 983	17	0.9	1037	5	AF044131	AF044131 Oncorhync
c 984	17	0.9	1037	5	AF044135	AF044135 Oncorhync
c 985	17	0.9	1037	5	AF044140	AF044140 Oncorhync
c 986	17	0.9	1037	5	AF044142	AF044142 Oncorhync
c 987	17	0.9	1037	5	AF044143	AF044143 Oncorhync
c 988	17	0.9	1037	5	AF044144	AF044144 Oncorhync
c 989	17	0.9	1037	5	AF044151	AF044151 Oncorhync
c 990	17	0.9	1037	5	AF044152	AF044152 Oncorhync
c 991	17	0.9	1037	5	AF044153	AF044153 Oncorhync
c 992	17	0.9	1037	5	AF044156	AF044156 Oncorhync
c 993	17	0.9	1037	5	AF044162	AF044162 Oncorhync
c 994	17	0.9	1037	5	AF044163	AF044163 Oncorhync
c 995	17	0.9	1037	5	AF044164	AF044164 Oncorhync
c 996	17	0.9	1038	5	AF044130	AF044130 Oncorhync
c 997	17	0.9	1038	5	AF044133	AF044133 Oncorhync
c 998	17	0.9	1038	5	AF044134	AF044134 Oncorhync
c 999	17	0.9	1038	5	AF044139	AF044139 Oncorhync
1000	17	0.9	1038	5	AF044148	AF044148 Oncorhync

ALIGNMENTS

RESULT	1					
LOCUS	AR120238	1999 bp	DNA	linear	PAT 16-MAY-2001	
DEFINITION	Sequence 1 from patent US 6159469.					
ACCESSION	AR120238					
VERSION	AR120238.1	GI:14103814				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1999)					
AUTHORS	Choi, G. H., Kunsch, C. A., Barash, S. C., Dillon, P. J., Dougherty, B., Fannon, M. R. and Rosen, C. A.					
TITLE	Streptococcus pneumoniae antigens and vaccines					
JOURNAL	Patent: US 6159469-A 12-DEC-2000;					
FEATURES	Location/Qualifiers					
source	1..1999					

QY 961 ACCGATCACAGACTATGCTCCTCGCTGGAGTACGGTGTCTAGCAATCAACTGCTACTAT 1020
 DB 961 ACCGATCACAGACTATGCTCCTCGCTGGAGTACGGTGTCTAGCAATCAACTGCTACTAT 1020
 QY 1021 CGTTCAGATGACCCCTATAACTACCTCGGCAAAATACCTCTGTTTATAACTGGGATAG 1080
 DB 1021 CGTTCAGATGACCCCTATAACTACCTCGGCAAAATACCTCTGTTTATAACTGGGATAG 1080
 QY 1081 GGGCTACTTTGGCAACATCACCTTTGCAATACGCCCTGCAACAAATCGCGAAACGTCGCCAGC 1140
 DB 1081 GGGCTACTTTGGCAACATCACCTTTGCAATACGCCCTGCAACAAATCGCGAAACGTCGCCAGC 1140
 QY 1141 CGTGGAACTCTAAACAAGTCGGACTCAACCGCGCAAGACTTTCTTAATGGTCTAGG 1200
 DB 1141 CGTGGAACTCTAAACAAGTCGGACTCAACCGCGCAAGACTTTCTTAATGGTCTAGG 1200
 QY 1201 AATCGACTACCCCAAGTATTCACTACTCAATGCCATTTCAAGTAAACACCAACGAAATCAGA 1260
 DB 1201 AATCGACTACCCCAAGTATTCACTACTCAATGCCATTTCAAGTAAACACCAACGAAATCAGA 1260
 QY 1261 CAAAAAATATGGAGCAAGTAGTGAAAGATGGCTGCTGCTTACGCTGCCTTTTCAAAATGG 1320
 DB 1261 CAAAAAATATGGAGCAAGTAGTGAAAGATGGCTGCTGCTTACGCTGCCTTTTCAAAATGG 1320
 QY 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTAGTGATGGGAGTGAATAA 1380
 DB 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTAGTGATGGGAGTGAATAA 1380
 QY 1381 AGAGTTCTTAATGTGCGAACTCGTGCCATGAAGGAACGACAGCCATATATGATGCCGA 1440
 DB 1381 AGAGTTCTTAATGTGCGAACTCGTGCCATGAAGGAACGACAGCCATATATGATGCCGA 1440
 QY 1441 CATGATGAACCAAGCTTGTGACTTATGGAATGGAGTGGCTGCTTACGCTGCTTGGCTGCC 1500
 DB 1441 CATGATGAACCAAGCTTGTGACTTATGGAATGGAGTGGCTGCTTACGCTGCTTGGCTGCC 1500
 QY 1501 TCAGGCTGTAAACAGGAACTCTAATCATACAGACGAGGAAATGGAACCAATCAATCA 1560
 DB 1501 TCAGGCTGTAAACAGGAACTCTAATCATACAGACGAGGAAATGGAACCAATCAATCA 1560
 QY 1561 GACCTCTCAATTTGTAGCACTGTATCAATTTTGTGCTATACGCTGAATATATCAAT 1620
 DB 1561 GACCTCTCAATTTGTAGCACTGTATCAATTTTGTGCTATACGCTGAATATATCAAT 1620
 QY 1621 GCGTGTATGACAGGCTATTCTAACCGTCTGACACCACTTTGAGGCAATGGCCTTACGGT 1680
 DB 1621 GCGTGTATGACAGGCTATTCTAACCGTCTGACACCACTTTGAGGCAATGGCCTTACGGT 1680
 QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAA 1740
 DB 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAA 1740
 QY 1741 TTGGAATATACAGAGGGCTCTACAGAAATGAGAAATTCGATTTTAAATGGTCTCG 1800
 DB 1741 TTGGAATATACAGAGGGCTCTACAGAAATGAGAAATTCGATTTTAAATGGTCTCG 1800
 QY 1801 TTTCTAGTGGAACTACCTGCTCCCAACAACCCCATCAACTGAAAGTTTCAAGCTTCATC 1860
 DB 1801 TTTCTAGTGGAACTACCTGCTCCCAACAACCCCATCAACTGAAAGTTTCAAGCTTCATC 1860
 QY 1861 ATCAGATAGTTCAACTTACAGTCTAGCTCAACCTCAAGCACAATAATAGTACGAC 1920
 DB 1861 ATCAGATAGTTCAACTTACAGTCTAGCTCAACCTCAAGCACAATAATAGTACGAC 1920
 QY 1921 TACCAATCTTAACATAATACGCAACAAATCAATACACCCCTGATCAACAAATCAGAA 1980
 DB 1921 TACCAATCTTAACATAATACGCAACAAATCAATACACCCCTGATCAACAAATCAGAA 1980
 QY 1981 TCCCTCAACCAAGCAACCA 1999
 DB 1981 TCCCTCAACCAAGCAACCA 1999

RESULT 2
 BD003825/c
 LOCUS
 DEFINITION Polynucleotide of Streptococcus pneumoniae and sequence.
 ACCESSION BD003825
 VERSION BD003825.1 GI:18631786
 KEYWORDS JP 2001501833-A/145.
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 10711)
 AUTHORS Kunsch,C.A., Choi,G.H., Dillon,P.J., Rosen,C.A., Bara,S.C., Fannon,M. and Dougherty,B.A.
 TITLE Polynucleotide of Streptococcus pneumoniae and sequence
 JOURNAL Patent: JP 2001501833-A 145 13-FEB-2001;
 HUMAN GENOME SCIENCES INC
 COMMENT OS Unidentified
 PN JP 2001501833-A/145
 PD 13-FEB-2001
 PF 30-OCT-1997 JP 1998520718
 PI CHARLES A KUNSCH,GIL H CHOI,PATRICK J DILLON,CRAIG A ROSEN, PI STEVEN C BARASH,
 PI MICHAEL FANNON,BRIAN A DOUGHERTY
 PC C12N15/09,A01K67/027,C07K14/315,C07K16/12,C12N1/15,C12N1/19,
 PC C12N1/21,
 PC C12N5/10,C12P21/02,C12Q1/68,G06F17/30,C12N15/00,C12N5/00, PC G06F15/40
 CC Strandedness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers
 FT source 1..10711 /organism='Unidentified'.
 FEATURES
 source Location/Qualifiers
 1..10711 /organism='unidentified'
 /db_xref='taxon:32644'
 BASE COUNT 3280 a 2056 c 2412 g 2963 t
 ORIGIN

Query Match 100.0%; Score 1999; DB 6; Length 10711;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAAATCTACGACAAATAAAATCAACTCATGTGCTGACTTGGGTTCTGAACGGCGCTCAA 60
 DB 9767 TAAATCTACGACAAATAAAATCAACTCATGTGCTGACTTGGGTTCTGAACGGCGCTCAA 9708
 QY 61 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
 DB 9707 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 9648
 QY 121 TCGCTTCTTCGACCACAGGGGATTTGATACCATCCCTATCTCGGGAGCTTCTTGGCGAA 180
 DB 9647 TCGCTTCTTCGACCACAGGGGATTTGATACCATCCCTATCTCGGGAGCTTCTTGGCGAA 9588
 QY 181 TCTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCACCACACAGTTGATTAAGTTGAC 240
 DB 9587 TCTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCACCACACAGTTGATTAAGTTGAC 9528
 QY 241 TTACTTTTCAACTTCCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
 DB 9527 TTACTTTTCAACTTCCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 9468
 QY 301 AGGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTCACTACTATATAATAA 360
 DB 9467 AGGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTCACTACTATATAATAA 9408
 QY 361 GGTCTACATGTCTAATGGGAACATATGGAATGCAGACAGCTCAAAACTACTATGGTAA 420
 DB 9407 GGTCTACATGTCTAATGGGAACATATGGAATGCAGACAGCTCAAAACTACTATGGTAA 9348

QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGGCTGGGTAATGCCCTCAGGCACC 480
|||||
Db 9347 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGGCTGGGTAATGCCCTCAGGCACC 9288
|||||
QY 481 AAACCAATATGACCCCTATTACATCCAGAACAGCCGCGGAACTTGGTCTTT 540
|||||
Db 9287 AAACCAATATGACCCCTATTACATCCAGAACAGCCGCGGAACTTGGTCTTT 9228
|||||
QY 541 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
|||||
Db 9227 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 9168
|||||
QY 601 ACCAATTACTGATGGACTACAAGTCTCAAAATCAGCAAGTAATATACCCTGCTTACATGA 660
|||||
Db 9167 ACCAATTACTGATGGACTACAAGTCTCAAAATCAGCAAGTAATATACCCTGCTTACATGA 9108
|||||
QY 661 TAATTACCTCAAGGAGTCAATCAATCAAGTTGAGGAAGAAACAGGCTATAACCTACTCAC 720
|||||
Db 9107 TAATTACCTCAAGGAGTCAATCAATCAAGTTGAGGAAGAAACAGGCTATAACCTACTCAC 9048
|||||
QY 721 AACTGGGATGGATGCTCACAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
|||||
Db 9047 AACTGGGATGGATGCTCACAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 8988
|||||
QY 781 TTACAAATACAGCAAGTAATGCTTGCCTATCCAGACGATGAATTCGAAGTGCCTTACCAT 840
|||||
Db 8987 TTACAAATACAGCAAGTAATGCTTGCCTATCCAGACGATGAATTCGAAGTGCCTTACCAT 8928
|||||
QY 841 TGTGTATGTTCTTAACGGTAAAGTCAATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
|||||
Db 8927 TGTGTATGTTCTTAACGGTAAAGTCAATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 8868
|||||
QY 901 TGTTCCTCTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGGGATCAACTATGAA 960
|||||
Db 8867 TGTTCCTCTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGGGATCAACTATGAA 8808
|||||
QY 961 ACGATACAGACTATGCTTCCTGCTTGGAGTACGGTGTCTAGGATCAACTGCTACTAT 1020
|||||
Db 8807 ACGATACAGACTATGCTTCCTGCTTGGAGTACGGTGTCTAGGATCAACTGCTACTAT 8748
|||||
QY 1021 CGTTACGATGACCCCTATAACTACCTCGGCAATACTCTGTTTATAACTGGGATAG 1080
|||||
Db 8747 CGTTACGATGACCCCTATAACTACCTCGGCAATACTCTGTTTATAACTGGGATAG 8688
|||||
QY 1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAATCGCGAAAGCTCCACG 1140
|||||
Db 8687 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAATCGCGAAAGCTCCACG 8628
|||||
QY 1141 CGTGGAACTCTAAACAAGTCCGACTCAACCGCGCAAGACTTTCTTAATGGTCTAGG 1200
|||||
Db 8627 CGTGGAACTCTAAACAAGTCCGACTCAACCGCGCAAGACTTTCTTAATGGTCTAGG 8568
|||||
QY 1201 AATCGACTACCAAGTATTCTACTCTCAATGCCATTTCAAGTAACACACCGCAATCAGA 1260
|||||
Db 8567 AATCGACTACCAAGTATTCTACTCTCAATGCCATTTCAAGTAACACACCGCAATCAGA 8508
|||||
QY 1261 CAAAAAATATGGAGCAAGTAGTGAAGAAGTGGCTGCTGCTTACGCTGCCCTTTGCAAAATGG 1320
|||||
Db 8507 CAAAAAATATGGAGCAAGTAGTGAAGAAGTGGCTGCTGCTTACGCTGCCCTTTGCAAAATGG 8448
|||||
QY 1321 TGGAACTTACTATAACCAATGTATATCCATAAGTCGCTTTAGTGATGGAGTGAANA 1380
|||||
Db 8447 TGGAACTTACTATAACCAATGTATATCCATAAGTCGCTTTAGTGATGGAGTGAANA 8388
|||||
QY 1381 AGAGTTCTCTAATGTCGGAACTCGGCCATGAAGGAACGACAGCTTATATGATGACCA 1440
|||||
Db 8387 AGAGTTCTCTAATGTCGGAACTCGGCCATGAAGGAACGACAGCTTATATGATGACCA 8328
|||||
QY 1441 CATGATGAACAGCTTTGACTTATGGAATGGAAGTGGAGCAATGCCTATCTTGTCTGGCTCCC 1500
|||||
Db 8327 CATGATGAACAGCTTTGACTTATGGAATGGAAGTGGAGCAATGCCTATCTTGTCTGGCTCCC 8268
|||||
QY 1501 TCAGCTGGTAAACAGGAACTCTAACTATACAGACGAGGAATTTGAAACCAACATCAA 1560

Db 8267 TCAGCTGGTAAACAGGAACTCTAACTATACAGACGAGGAATTTGAAACCAACATCAA 8208
|||||
QY 1561 GACCTCTCAATTTGTAGCACCTTGATGAACCTATTTCCTGGCTATACGCGTAAATATTCAAT 1620
|||||
Db 8207 GACCTCTCAATTTGTAGCACCTTGATGAACCTATTTCCTGGCTATACGCGTAAATATTCAAT 8148
|||||
QY 1621 GGCTGATGACGAGGCTATTCTAACCGCTCTGACACACCTTTGAGGCAATGGCCTTACGGT 1680
|||||
Db 8147 GGCTGATGACGAGGCTATTCTAACCGCTCTGACACACCTTTGAGGCAATGGCCTTACGGT 8088
|||||
QY 1681 CGCTGCCAAAGTTTACCCTCTATGATGACCTTACCTGCTGAGGAAAGCAATCCAGAA 1740
|||||
Db 8087 CGCTGCCAAAGTTTACCCTCTATGATGACCTTACCTGCTGAGGAAAGCAATCCAGAA 8028
|||||
QY 1741 TTGGAAATATACGAGGGGCTCTACAGAAATGGAGAAATTCGTATTAAAAAATGGTCTCG 1800
|||||
Db 8027 TTGGAAATATACGAGGGGCTCTACAGAAATGGAGAAATTCGTATTAAAAAATGGTCTCG 7968
|||||
QY 1801 TTCTAGTGGAACTACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
|||||
Db 7967 TTCTAGTGGAACTACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 7908
|||||
QY 1861 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCAACGACACAAATAATAGTAGC 1920
|||||
Db 7907 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCAACGACACAAATAATAGTAGC 7848
|||||
QY 1921 TACCAATCTCAATAATACGCAACAATCAATAACAACCCCTGATCAACAAATCAGAA 1980
|||||
Db 7847 TACCAATCTCAATAATACGCAACAATCAATAACAACCCCTGATCAACAAATCAGAA 7788
|||||
QY 1981 TCCTCAACGACGACCAACCA 1999
|||||
Db 7787 TCCTCAACGACGACCAACCA 7769
|||||
RESULT 3
AE007349/c
LOCUS
DEFINITION Streptococcus pneumoniae TIGR4 section 32 of 194 of the complete genome. BCT 31-AUG-2001
ACCESSION AE007349 AE005672
VERSION AE007349.1 GI:14971837
KEYWORDS
SOURCE Streptococcus pneumoniae TIGR4
ORGANISM Streptococcus pneumoniae TIGR4
REFERENCE 1 (bases 1 to 11334)
AUTHORS Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D., Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J., Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D., Umayam, L.A., White, O., Salzberg, S.L., Lewis, M.R., Radune, D., Holtzapple, E., Khouri, H., Wolf, A.M., Utterback, T.R., Hansen, C.L., McDonald, L.A., Feldblyum, T.V., Angiuoli, S., Dickson, F., Hickey, E.K., Holt, I.E., Loftus, B.J., Yang, F., Smith, H.O., Venter, J.C., Dougherty, B.A., Morrison, D.A., Hollingshead, S.K. and Fraser, C.M.
TITLE Complete genome sequence of a virulent isolate of Streptococcus pneumoniae
JOURNAL Science. 293 (5529), 498-506 (2001)
MEDLINE 21357209
PUBMED 11463916
REFERENCE 2 (bases 1 to 11334)
AUTHORS Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D., Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J., Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D., Umayam, L.A., White, O., Lewis, M.R., Radune, D., Holtzapple, E., Khouri, H., Wolf, A.M., Utterback, T.R., Hansen, C.L., McDonald, L.A., Feldblyum, T.V., Angiuoli, S., Gesuwan, P., Hickey, E.K., Holt, I.E., Loftus, B.J., Ujwal, M.L., Yang, F., Smith, H.O., Venter, J.C., Dougherty, B.A., Morrison, D.A., Hollingshead, S.K. and Fraser, C.M.
TITLE Direct Submission

JOURNAL	Submitted (29-JUN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES	Location/Qualifiers
source	1. .11334
	/organism="Streptococcus pneumoniae TIGR4"
	/strain="TIGR4"
	/db_xref="taxon:170187"
gene	complement(114. .2273)
	/gene="SP0369"
CDS	complement(114. .2273)
	/gene="SP0369"
	/codon_start=1
	/transl_table=11
	/product="penicillin-binding protein 1A"
	/protein_id="AAK74536.1"
	/db_xref="GI:14971838"
	/translation="MNKPTILRLIKYLSIFLSILVIAAIVLGGGVFFVYVSKAPSLSE SKLVATTSKIIDYDNKQLIADLGSERRVNAQNDIPTDLVKAIVSIEDHRFFDRGID TIRILGAFRLNLSQSGSTLQQLIKLITFTSTSDQTSIRKAQAWLAIALEQK ATKQELIYYINKVYMSNGNYGMOTAAQYIKGDLNNLSLQALLAGMPAQNPQDP YSHPEAQDRNLVLESENKQGYISAEQYERAVNTPITDGLQSLKASNYPAYMDNYL KEVINQVEETGYNLLTTGMDVYINVDQEAQHLDWIYTDGLQSLVAPDQELQVASTIV DVSKGVIQAOLGARHOSNVSEGINQAVETNRDMGSTMKPIITDVAPELEYGYDSTAT IVHDEPNVPCNTPNVYNKRGYEGNITLOALQOSRVNPAVELNKYGLNRAKTEFLN GLGIDYPSIHYSNALSNTTESDKKYGASSEKMAAAYAAFAANGTYIYPMTHKVVFS DGSKEFNSVGTWRAMKETATYMMTDMKTVITGTGRNAYLAWLPQAGKTGTSTNDE BIENHIKTSQVAPDELFAQYTRYKSMVLTGYSNRLTPLVNGLTVAAKYVRSMTY LSEGSNPDNMNIPGLYRNGEFVKNGARSTWNSPAPQPPSTESSSSSSSSTSSQSS STPTSTNNSTTTPNNNTTQQSNTTPDQGNQNPQPAQ"
gene	complement(2270. .2866)
	/gene="SP0370"
CDS	complement(2270. .2866)
	/gene="SP0370"
	/codon_start=1
	/transl_table=11
	/product="recombination protein U"
	/protein_id="AAK74537.1"
	/db_xref="GI:14971839"
	/translation="MVNYPHKVSSQKROTSISQPKNFANRGMSEFKMINATNDYYLSQ GLAVIHKKPTQIQIVDYPQSRRAKIVAEYFRQASTDYSGVNGYIDFEVKETQ KRAIPMKNFHQIQHMEQVLAQQICFVLLHFSSQETYLPLPDLIRFVHQDGQK SMPLEYIREGYEIKAGAFPOIPVLNVIKEHLGGKTR"
gene	2932. .3459
	/gene="SP0371"
CDS	2932. .3459
	/gene="SP0371"
	/note="identified by Glimmer2; putative"
	/codon_start=1
	/transl_table=11
	/product="conserved hypothetical protein"
	/protein_id="AAK74538.1"
	/db_xref="GI:14971840"
	/translation="MATALVLYGSFAFLGFSDDKPRLLKKAIRKDLKLEMAADGVS WLVTGSLGFEVWVLEVAQEMKTEVGQLATIFAFETHGENNENGNOMKLSRFQVDF KYAYAPREHYGQLRDYQQFLLENTTSYLYFDENETKLAYFYQKKMKNQDYFKRL TDELNELAENFSEN"
gene	3529. .3858
	/gene="SP0372"
CDS	3529. .3858
	/gene="SP0372"
	/note="identified by Glimmer2; putative"
	/codon_start=1
	/transl_table=11
	/product="conserved hypothetical protein"
	/protein_id="AAK74539.1"
	/db_xref="GI:14971841"
	/translation="WASLIIFSADKIFEQEFGREVRGYNKVEVDFLDVYIKDYETAA LVKSLRQIADLKBELTRKPKPSPVQAPLEAAITSSMTNFDILRLRLRLEKEVFGKQ ILDNSDF"
gene	4344. .5501
	/gene="SP0373"
CDS	4344. .5501
	/gene="SP0373"

	/note="identified by Glimmer2; putative"
	/codon_start=1
	/transl_table=11
	/product="conserved hypothetical protein"
	/protein_id="AAK74540.1"
	/db_xref="GI:14971842"
	/translation="MKKEFNLIATVAAGLEAVVGREVRELGYDQOVENGVRVRFQGDVR AIIETNMLRAADRIKIIIVGTFFPAKTFEELFOGVFALDWMENYLPILGARFPPISKACVK SKLHNPSVQAISSKAVKIKKQHYARPEGVPMENGEPEFKIEVILKDKVATVMTDIT GSSLFKRGVYRTKGAIPKENMAAIILOLSNWPDKPLIDPTGSGGTFCIEAVMTARK MAPGLRRSFAFEENWNI SDRLIQEVRTAAKKVDRELELDIMGCDIDARMVEIAKANA QVAGVAGDITFKOMRVQDLRSKINGVIISNPYGERISDDAGYTKLVAEMGQVAPL KTWSKFILTSDEAFESKYSQADKKRKLNGTLKVDLYQIFGQVRKROEVK"
gene	5514. .6908
	/gene="SP0374"
CDS	5514. .6908
	/gene="SP0374"
	/note="identified by Glimmer2; putative"
	/codon_start=1
	/transl_table=11
	/product="hypothetical protein"
	/protein_id="AAK74541.1"
	/db_xref="GI:14971843"
	/translation="MSKRRNRHKKGEQEPQDFDEAKELTVQGAIRKNEEVESVLP EDSIDLVKQHRDEIADQYIKKEEFVETQSDLDLOEMREAVEKSSASSEEV PSESDILPLDDEEGQLDPLLDENPTTEVEEQNLSRLDQEDSEKSKKGF ILTVLALVSVIICVAYVYQVARSKEIETQSQTANSDVDNFNTLYDAFYDTSN KTAIKNQKDLSQLTKLLDKLEGSREHTLAKSKYDSLATQIKAIQDVNAQPEKPAIV DGVLTKAKAKSDAKFTDKTGTNLDKDAISLGKSOOTSTSSSSSOPTSSSSSS QASNTSEPKPSSNSETSSSEVNMGLSSAGAVORSASRAYNQSALDDNSNAW PADGVLSQILATSRSGYITGDQYILERVNIVNGYNYLKYKPGTYLFTLCKTGY FVNGAGHADDLDY"
gene	6984. .8408
	/gene="SP0375"
CDS	6984. .8408
	/gene="SP0375"
	/note="identified by match to TIGR protein family HMM TIGR00873"
	/codon_start=1
	/transl_table=11
	/product="6-phosphogluconate dehydrogenase, decarboxylating"
	/protein_id="AAK74542.1"
	/db_xref="GI:14971844"
	/translation="MTKANFGVYGVMAVMGRNLALNIESRGYTVATYINRSKEDVIA CHPEKFPVSYDVSFVNSIEKPRRIMLMVQAGPTDATIQALLPHLDKGDILDGGN TFYKDTIRRENLANSINFGTGYSGEGKALEGFSIMPGQKAYELVADVLEIS AKAPEDGKPCVTYIGPDGAGHVYKMHNGIEYQDMOLIAESYDLMOHLGLLSAEDMAE IFTENWKGELSYLLEITADILSRKDDQDGPVDYILDAAGNKGTKGWTSSQLDL GVPLSLITSFVARYISTYKEERVHASKVLPAPAFNFGDKAELEIKRQALYFSKI ISYAQFQALRVASKENNNLPPADIASIWRDGCIIIRSFLOKITDAYNRDADLANLL LDEYFLDTAKYQQAQVDRDIALAVQAGVPVPTFSAAITFYDSYRSADLPANLIQAQRD YFGAITYQKKKEGTFHYSWYDEK"
gene	8420. .9109
	/gene="SP0376"
CDS	8420. .9109
	/gene="SP0376"
	/codon_start=1
	/transl_table=11
	/product="DNA-binding response regulator"
	/protein_id="AAK74543.1"
	/db_xref="GI:14971845"
	/translation="MGRKILLILEKERNLAHFLSLQEQRYVDVLEEGQKALSMAQ TDYDLMLNVLNGLGDMMAQDFAEKRSRTPASVIMILDHWDQLEQVYVQFAVSYIY KPVLIENLVARISAFIRGRDFTDQHCILMKVPTYRNLRIQVDEHHTVYRGEMIALTR REYDLATLMGSKKVLTRQELLLESVMKYESATETNIDVYIRYLSKLDLVKGQKSYIK TVRGYITMQE"
gene	9208. .10230
	/gene="SP0377"
CDS	9208. .10230
	/gene="SP0377"
	/codon_start=1
	/transl_table=11

/product="choline binding protein C"									
/protein_id="AAK74544.1"									
/db_xref="GI:14971847"									
/translation="MKLLKKMMQVALAVFFGLLATNTVFANTTGGRFVVDKDNKYYV									
KDDHKAITWHKIDGTYIFDGIEMVWQYLEIPGTGRDNLFPQNPVNEIGLQEKW									
YFGDGALLSQTDKQVLEAKTSENTGVYGEQYPLSAEKRTYYFDNNYAVKTYIWE									
BDHWYLLNKLNGFSDSYNPLPIGEVAKGWTQDPHVTIDIDRSKPAPWYLDASGKML									
TDMQVNGKWTYFSGSGSMATGWKVRGKWTYLDNKGMDKMTGWOLGNKWYILRSSG									
AMVTCWQDGSWYLYLDPNSGDMKLGWTKVNGKWTYLLNSNGAMVTSQTDIDKVVNFA									
SSGEWI"									
10248..11246									
/gene="sp0378"									
10248..11246									
/gene="sp0378"									
/codon_start=1									
/transl_table=11									
/product="choline binding protein J"									
/protein_id="AAK74545.1"									
/db_xref="GI:14971847"									
Query Match 100.0%; Score 1999; DB 1; Length 11334;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	TAAATCTACGACATATAAATCAACTCATGCTGACTTGGGTCTGAAACGGCGGCTCAA	60						
DB	2115	TAAATCTACGACATATAAATCAACTCATGCTGACTTGGGTCTGAAACGGCGGCTCAA	2056						
QY	61	TGCCAAGCTAATGATATTCACACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA	120						
DB	2055	TGCCAAGCTAATGATATTCACACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA	1996						
QY	121	TGCTTCTTCGACACAGGGGATTCATACCATCCGTATCTCGGGAGCTTCTTGCGCAA	180						
DB	1995	TGCTTCTTCGACACAGGGGATTCATACCATCCGTATCTCGGGAGCTTCTTGCGCAA	1936						
QY	181	TTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCACCCCAACAGTTGATTAAAGTTGAC	240						
DB	1935	TCTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCACCCCAACAGTTGATTAAAGTTGAC	1876						
QY	241	TTACTTTTCAACTTCGACTTCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	300						
DB	1875	TTACTTTTCAACTTCGACTTCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	1816						
QY	301	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	360						
DB	1815	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	1756						
QY	361	GGCTACATGCTAATGGGAATATGGAATGCAGACAGAGCTCAAAACTACTATGGTAA	420						
DB	1755	GGCTACATGCTAATGGGAATATGGAATGCAGACAGAGCTCAAAACTACTATGGTAA	1696						
QY	421	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGCTGGCTGGAATGCCTCAGGCACC	480						
DB	1695	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGCTGGCTGGAATGCCTCAGGCACC	1636						
QY	481	AAACCAATATGACCCCTATTACATCCAGAACGAGCCCAAGACCGCCGAAACTTGGCTCT	540						
DB	1635	AAACCAATATGACCCCTATTACATCCAGAACGAGCCCAAGACCGCCGAAACTTGGCTCT	1576						
QY	541	ATCTGAAATGAANAATCAAGGCTAGATCTGCTGAACAGTATGAGAAAGCAGTCAATAC	600						
DB	1575	ATCTGAAATGAANAATCAAGGCTAGATCTGCTGAACAGTATGAGAAAGCAGTCAATAC	1516						
QY	601	ACCAATCTAGTGGACTACAAGTCTCAATTCAGCAAGTAATACCCTGCTTACATGA	660						
DB	1515	ACCAATCTAGTGGACTACAAGTCTCAATTCAGCAAGTAATACCCTGCTTACATGA	1456						
QY	661	TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGAAACAGGCTATAAACCCTACTAC	720						
DB	1455	TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGAAACAGGCTATAAACCCTACTAC	1396						
QY	721	AACTGGGATGGATGCTACACAAATGTAGACCAAGAGCTCAAAAACATCTCTGGGATAT	780						

DB	1395	AACTGGGATGGATGCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	1336						
QY	781	TTACAATACAGACGAATACGTTCCCTATTCAGACGATGAATTCGAAGTCGCTTCTACCAT	840						
DB	1335	TTACAATACAGACGAATACGTTCCCTATTCAGACGATGAATTCGAAGTCGCTTCTACCAT	1276						
QY	841	TGTTGATGTTCTTAACGGTAAAGTCATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA	900						
DB	1275	TGTTGATGTTCTTAACGGTAAAGTCATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA	1216						
QY	901	TGTTTCTCTCGGAATTAACCAAGCAGTAGAACAACACCGCAGCTGGGGATCAACTATGAA	960						
DB	1215	TGTTTCTCTCGGAATTAACCAAGCAGTAGAACAACACCGCAGCTGGGGATCAACTATGAA	1156						
QY	961	ACCGATCACAGACTATGCTCCCTTGGAGTAGGTGTCTAGGATCAACTGCTACTAT	1020						
DB	1155	ACCGATCACAGACTATGCTCCCTTGGAGTAGGTGTCTAGGATCAACTGCTACTAT	1096						
QY	1021	CGTTCACGATGAGCCCTATAAATACCTGAGCAAAATACCTCTGTTTATAAATCGGATAG	1080						
DB	1095	CGTTCACGATGAGCCCTATAAATACCTGAGCAAAATACCTCTGTTTATAAATCGGATAG	1036						
QY	1081	GGCTACTTTGGCAACATCACCTTGCATACGCCCTGCAACAATCGGAAACGTCGCCAGC	1140						
DB	1035	GGCTACTTTGGCAACATCACCTTGCATACGCCCTGCAACAATCGGAAACGTCGCCAGC	976						
QY	1141	CGTGGAACTCTAAACAAGTGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG	1200						
DB	975	CGTGGAACTCTAAACAAGTGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG	916						
QY	1201	AATCGACTACCCAAAGTATTCACTACTCAATGCCATTTCAAGTAACACACCAATCAGA	1260						
DB	915	AATCGACTACCCAAAGTATTCACTACTCAATGCCATTTCAAGTAACACACCAATCAGA	856						
QY	1261	CAAAAAATATGGAGCAAGTAGTGAAAAGATGGCTGTGTGTAGCTGTCCTTTGCAAAATGG	1320						
DB	855	CAAAAAATATGGAGCAAGTAGTGAAAAGATGGCTGTGTGTAGCTGTCCTTTGCAAAATGG	796						
QY	1321	TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGTCTTTAGTGTAGGGAGTGA	1380						
DB	795	TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGTCTTTAGTGTAGGGAGTGA	736						
QY	1381	AGAGTTCTCTAATGTGCGAACTCGTCCCATGAAGGAAAGCAGACGCTATATGATGACCGA	1440						
DB	735	AGAGTTCTCTAATGTGCGAACTCGTCCCATGAAGGAAAGCAGACGCTATATGATGACCGA	676						
QY	1441	CATGATGAAACAGCTTGTGACTTATGGAATGGACGAAATGCCTATCTTGTGCTGCC	1500						
DB	675	CATGATGAAACAGCTTGTGACTTATGGAATGGACGAAATGCCTATCTTGTGCTGCC	616						
QY	1501	TCAGGCTGTAAACAGGAACCTCTAATCTACAGACGAGGAATTTGAAACCCACATCAA	1560						
DB	615	TCAGGCTGTAAACAGGAACCTCTAATCTACAGACGAGGAATTTGAAACCCACATCAA	556						
QY	1561	GACCTCTCAATTTGTAGCAGCTGATGAACCTATTTGCTGGCTATACGCGTAAATATTTCAAT	1620						
DB	555	GACCTCTCAATTTGTAGCAGCTGATGAACCTATTTGCTGGCTATACGCGTAAATATTTCAAT	496						
QY	1621	GGCTGATGACAGGCTATTTAAACCGCTGACACCACTTGTGTAGGCAATGGGCTTACGGT	1680						
DB	495	GGCTGATGACAGGCTATTTAAACCGCTGACACCACTTGTGTAGGCAATGGGCTTACGGT	436						
QY	1681	CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAGGAGCAATCCCAAGA	1740						
DB	435	CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAGGAGCAATCCCAAGA	376						
QY	1741	TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGTAATTTAAAAATGGTCTCG	1800						
DB	375	TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGTAATTTAAAAATGGTCTCG	316						
QY	1801	TTCTAGGTGGAATCACTGCTGCCACACAAACCCCATCAACTGAAAGTTTCAAGCTCATC	1860						

Db 1179 CGTTCACGATGACCCCTATATATACCTCGGACAAATACTCCTGTTTATAACTGGGATAG 1238
 QY 1081 GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACAATCGCGAAGCTCCAGC 1140
 Db 1239 GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACAATCGCGAAGCTCCAGC 1298
 QY 1141 CGTGGAACTCTAAACAAGGTCGGACTCAACCGCGCAAGACTTTCTCTAAATGGTCTAGG 1200
 Db 1299 CGTGGAACTCTAAACAAGGTCGGACTCAACCGCGCAAGACTTTCTCTAAATGGTCTAGG 1358
 QY 1201 AATGACTACCAAGTATTCACCTACTCAATGCCATTTCAAGTAACACACCAATCAGA 1260
 Db 1359 AATGACTACCAAGTATTCACCTACTCAATGCCATTTCAAGTAACACACCAATCAGA 1418
 QY 1261 CAAAAAATATGAGCAAGTAGTGAAGAAGTGGCTGCTTACGCTGCTTTGCAAAATGG 1320
 Db 1419 CAAAAAATATGAGCAAGTAGTGAAGAAGTGGCTGCTTACGCTGCTTTGCAAAATGG 1478
 QY 1321 TGGAACTTACTATAACCAATGTATATCCATAAAGTCGCTTTAGTGATGGAGTGAAAA 1380
 Db 1479 TGGAACTTACTATAACCAATGTATATCCATAAAGTCGCTTTAGTGATGGAGTGAAAA 1538
 QY 1381 AGAGTCTCTAATGTCGGAACCTGTCGCCATGAAGGAACGACAGCCTATATGATACCGA 1440
 Db 1539 AGAGTCTCTAATGTCGGAACCTGTCGCCATGAAGGAACGACAGCCTATATGATACCGA 1598
 QY 1441 CATGATGAAACAGTCTTGACTTATGGAACCTGACGAAATGCTATCTTGTGCTGCC 1500
 Db 1599 CATGATGAAACAGTCTTGACTTATGGAACCTGACGAAATGCTATCTTGTGCTGCC 1658
 QY 1501 TCAGGCTGGTAAACAGGAACCTCTAATATACAGACGAGGAATTTGAAACACCATCAA 1560
 Db 1659 TCAGGCTGGTAAACAGGAACCTCTAATATACAGACGAGGAATTTGAAACACCATCAA 1718
 QY 1561 GACCTCTCAATTTGAGCAACCTGATGAATTTGCTGGCTATACGCGTAATATTCAT 1620
 Db 1719 GACCTCTCAATTTGAGCAACCTGATGAATTTGCTGGCTATACGCGTAATATTCAT 1778
 QY 1621 GCGTGTATGACAGGCTATTTCAACGCTCTGACACCACTTTGAGCAATGGCCTTACGGT 1680
 Db 1779 GCGTGTATGACAGGCTATTTCAACGCTCTGACACCACTTTGAGCAATGGCCTTACGGT 1838
 QY 1681 CCGTCCCAAGTTTACCGCTCTATGATGACCTACCTGCTGAGGAAGCAATCCAGAGA 1740
 Db 1839 CCGTCCCAAGTTTACCGCTCTATGATGACCTACCTGCTGAGGAAGCAATCCAGAGA 1898
 QY 1741 TTGGAATATACCAAGGGGCTCTACAGAAATGAGAATTCGTATTTAAAAATGGTCTCG 1800
 Db 1899 TTGGAATATACCAAGGGGCTCTACAGAAATGAGAATTCGTATTTAAAAATGGTCTCG 1958
 QY 1801 TTCTACGTGGAACCTCACCTGCTCCACAAACCCCATCACTGAAAGTTCAAGCTCATC 1860
 Db 1959 TTCTACGTGGAACCTCACCTGCTCCACAAACCCCATCACTGAAAGTTCAAGCTCATC 2018
 QY 1861 ATCAGATAGTTCACTTACAGCTAGCTCAACCACTCCAGCAACAATAATAGTACGAC 1920
 Db 2019 ATCAGATAGTTCACTTACAGCTAGCTCAACCACTCCAGCAACAATAATAGTACGAC 2078
 QY 1921 TACCAATCCTACATAATACCAACAATCAATACACCCCTGATCAACAAATCAGAA 1980
 Db 2079 TACCAATCCTACATAATACCAACAATCAATACACCCCTGATCAACAAATCAGAA 2138
 QY 1981 TCCTCAACACGACCAACCA 1999
 Db 2139 TCCTCAACACGACCAACCA 2157

RESULT 5
 LOCUS SPON8
 DEFINITION S. pneumoniae (45607) ponA gene for penicillin-binding protein 1a.
 ACCESSION X67873
 VERSION X67873.1 GI:47419

penicillin-binding protein 1a; ponA gene.
 Streptococcus pneumoniae.
 Streptococcus pneumoniae
 Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 Streptococcus.
 1 (bases 1 to 2157)
 Hakenbeck, R.
 Direct Submission
 Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek.
 Genetik, thnestr. 73, 1000 Berlin 33, FRG
 2 (bases 1 to 2157)
 Martin, C., Sibold, C. and Hakenbeck, R.
 Relatedness of penicillin-binding protein 1a genes from different
 clones of penicillin-resistant Streptococcus pneumoniae isolated in
 South Africa and Spain
 EMBO J 11 (11), 3831-3836 (1992)
 93010977 Location/Qualifiers
 1..2157
 /organism="Streptococcus pneumoniae"
 /strain="45607"
 /db_xref="taxon:1313"
 /clone="pcMK1-pcMK16"
 1..2157
 /gene="ponA"
 1..>2157
 /gene="ponA"
 /codon_start=1
 /transl_table=11
 /product="penicillin-binding protein 1a"
 /protein_id="CAA48073.1"
 /db_xref="GI:47420"
 /db_xref="SWISS-PROT:Q04707"
 SKLVATSSKIYDNKNQIADLGSERRVNAQNDIPTDLVAIVSIEDHRRFPDHRGID
 TIRLIGAFRLNLSQNSGSLTQQLIKLTFTSTSDTISRKAQGAWPAIQLEQK
 ATKEILTYINKVNSNGNYGMQTAQNYGKDLNLSLPQLALLAGMPAPNQYDP
 YSHPAAQDRRLNVLSENKQYISAEQYKAVNPTITDGLSKASNYPAYMDNYL
 KEVINQVEETGYNLLTTGMDVYFNVDQEAQKHLMDIYNTDEYVAYPDELYQVSTIV
 DVSNGKYIAQLGARHOSNVSFGINQAVETNRDWSGTMKPTIDVAPALEYQVSTIV
 IVHDEPNYPCNTNPVYNDRGYFNITLOVALQOSNRVPAVETLNKVLNRAKTELN
 GLGIDYPSIHYNSALISNTTESDKKYGASSERMAAAYAFANGGTYIYKPMIHKVYFS
 DGSSEFNSVTRAMKETATYMTTMMKTVLTITGTRNAYLAWLPQAQKTGTSNTDE
 EIEHNIKTQSFVADELEFAGYTRKYSMAVWTGYSNRLTPLYNGLTVAAKYVRSMTY
 LSEGSNPEDWNIPEGLYRNGEFVFKNGARSTWSSPAQPPSTESSSSSSSSSTSQSS
 STTPTSNNTTNNNTNQSNTPDQONQNPQAP"
 BASE COUNT 696 a 520 c 411 g 530 t
 ORIGIN
 Query Match 79.6%; Score 1591; DB 1; Length 2157;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1991; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 TAAATCTACGACATAAAAAATCAACTCATTCGCTGCTGTTCTGAAACGGCGTCAA 60
 Db 159 TAAATCTACGACATAAAAAATCAACTCATTCGCTGCTGTTCTGAAACGGCGTCAA 218
 QY 61 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGGCAATCGTTCTATCGAAGCA 120
 Db 219 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGGCAATCGTTCTATCGAAGCA 278
 QY 121 TCGCTTCTTCGACACAGGGGATTTGATACCATCGTATCTCTGGAGCTTTCTTCGCGAA 180
 Db 279 TCGCTTCTTCGACACAGGGGATTTGATACCATCGTATCTCTGGAGCTTTCTTCGCGAA 338
 QY 181 TCTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCAACCAAGCTTGATTAGTTGAC 240
 Db 339 TCTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCAACCAAGCTTGATTAGTTGAC 398
 QY 241 TTACTTTTCAACTTCGACTTCGACCAAGCTATTTCTCGTAGGCTCAGGAGCTTGGTT 300
 Db 399 TTACTTTTCAACTTCGACTTCGACCAAGCTATTTCTCGTAGGCTCAGGAGCTTGGTT 458

QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA 360
Db 459 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA 518
QY 361 GGTCTACATGCTTAATGGGAACATAGGAATGCAGACAGCAGCTCAAACTACTATGGTAA 420
Db 519 GGTCTACATGCTTAATGGGAACATAGGAATGCAGACAGCAGCTCAAACTACTATGGTAA 578
QY 421 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTTGGCTGGGAATGCCCTCAGCACCC 480
Db 579 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTTGGCTGGGAATGCCCTCAGCACCC 638
QY 481 AAACCAATATGACCCCTATTTCATATCCAGAAGCAGCCCAAGCAGCCGCGAACTTGGCTCT 540
Db 639 AAACCAATATGACCCCTATTTCATATCCAGAAGCAGCCCAAGCAGCCGCGAACTTGGCTCT 698
QY 541 ATCTGNAATGAANAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 699 ATCTGNAATGAANAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 758
QY 601 ACCAATTTACTGATGGACTACAAGTCTCAAAATCAGCAAGTAAATACCCCTGCTTACATGGA 660
Db 759 ACCAATTTACTGATGGACTACAAGTCTCAAAATCAGCAAGTAAATACCCCTGCTTACATGGA 818
QY 661 TAATTACTCTAAGGAAGTCAATCAATCAAGTTGAAGAAGAAACAGGCTATAAOCCTACTCAC 720
Db 819 TAATTACTCTAAGGAAGTCAATCAATCAAGTTGAAGAAGAAACAGGCTATAAOCCTACTCAC 878
QY 721 AACTGGGATGGATGCTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATAT 780
Db 879 AACTGGGATGGATGCTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATAT 938
QY 781 TTACAATACAGACGATACGTTGCCCTATCCAGACGATGAATGCAAGTGCCTTCTACCAT 840
Db 939 TTACAATACAGACGATACGTTGCCCTATCCAGACGATGAATGCAAGTGCCTTCTACCAT 998
QY 841 TGTGATGTTTCAAGGTAAGTCAATGCCCAGCTTAGGAGCAGCCATCAGTCAAGTAA 900
Db 999 TGTGATGTTTCAAGGTAAGTCAATGCCCAGCTTAGGAGCAGCCATCAGTCAAGTAA 1058
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGGATCAACTATGAA 960
Db 1059 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGGATCAACTATGAA 1118
QY 961 ACCGATCACAGACTATGCTTCCTGCTTGGAGTACGGTGTACGATTCAACTGCTACTAT 1020
Db 1119 ACCGATCACAGACTATGCTTCCTGCTTGGAGTACGGTGTACGATTCAACTGCTACTAT 1178
QY 1021 CGTTCAGATGAGCCCTATACTACCCCTGGGACAAATACCTGTTTATACTGGGATAG 1080
Db 1179 CGTTCAGATGAGCCCTATACTACCCCTGGGACAAATACCCCTGTTTATACTGGGATAG 1238
QY 1081 GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACATCGCAACGTCGAAAGCTCCAGC 1140
Db 1239 GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACATCGCAACGTCGAAAGCTCCAGC 1298
QY 1141 CGTGGAACTCTAAACAGGTCGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG 1200
Db 1299 CGTGGAACTCTAAACAGGTCGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTCGG 1358
QY 1201 AATGACTACCCCAAGTATTCATCTACTCAATGCGATTTCAAGTAACACACCGAATCAGA 1260
Db 1359 AATGACTACCCCAAGTATTCATCTACTCAATGCGATTTCAAGTAACACACCGAATCAGA 1418
QY 1261 CAAAAAATATGAGCAAGTAGTGAAGAATGCTGCTGCTTACGCTGCCCTTTGCAAAATGG 1320
Db 1419 CAAAAAATATGAGCAAGTAGTGAAGAATGCTGCTGCTTACGCTGCCCTTTGCAAAATGG 1478
QY 1321 TGGAACTTACTATAACCAATGTATATCCATAAAGTCGCTTTAGTGATGGGAGTCAAAA 1380
Db 1479 TGGAACTTACTATAACCAATGTATATCCATAAAGTCGCTTTAGTGATGGGAGTCAAAA 1538

QY 1381 AGAGTTTCTTAATGTGCGAACTCGTGCCATGAAGGAAACGACAGCCTATATATGATGACCGA 1440
Db 1539 AGAGTTTCTTAATGTGCGAACTCGTGCCATGAAGGAAACGACAGCCTATATATGATGACCGA 1598
QY 1441 CATGATGAACACAGTCTTGACTTATGAACTGGACGAATGCCTATCTTCTGCTTGGCTCCC 1500
Db 1599 AATGATGAACACAGTCTTGACTTATGAACTGGACGAATGCCTATCTTCTGCTTGGCTCCC 1558
QY 1501 TCAGGCTGTTAAAAAGGAACTCTAACTATACAGACGAGGAAATGAAACACACATCAA 1560
Db 1659 TCAGGCTGTTAAAAAGGAACTCTAACTATACAGATGAGAAATGAAACACACATCAA 1718
QY 1561 GACCTCTCAATTTGTAGACCTGATGAACATTTTGTGGCTATACGCGTAAATATTCAAT 1620
Db 1719 GACCTCTCAATTTGTAGACCTGATGAACATTTTGTGGCTATACGCGTAAATATTCAAT 1778
QY 1621 GGTGTATGGACAGGCTATTCTAACCGCTCTGACACCACTTGTAGGCAATGGCCTTACGGT 1680
Db 1779 GGTGTATGGACAGGCTATTCTAACCGCTCTGACACCACTTGTAGGCAATGGCCTTACGGT 1838
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTCTCTGAAGGAAGCAATCCAGAA 1740
Db 1839 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTCTCTGAAGGAAGCAATCCAGAA 1898
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTGTCTG 1800
Db 1899 TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTGTCTG 1958
QY 1801 TTCTAGTGAACCTACCTGCTCCACAACACCCCATCAACTCAAGTTTCAAGCTCATC 1860
Db 1959 TTCTAGTGAAGCTACCTGCTCCACAACACCCCATCAACTCAAGTTTCAAGCTCATC 2018
QY 1861 ATCAGATAGTTCAACTTTCACAGTCTAGCTCAACACTTCCAAAGCACAATAATAGTACGAC 1920
Db 2019 ATCAGATAGTTCAACTTTCACAGTCTAGCTCAACACTTCCAAAGCACAATAATAGTACGAC 2078
QY 1921 TACCAATCTTAACAATAATAGGCAACAATAACAACCCCTGATCAACAATAATCAGAA 1980
Db 2079 TACCAATCTTAACAATAATAGGCAACAATAACAACCCCTGATCAACAATAATCAGAA 2138
QY 1981 TCCTCAACCAAGCACAACCA 1999
Db 2139 TCCTCAACCAAGCACAACCA 2157
RESULT 6
AF210746
LOCUS
DEFINITION
AF210746
ACCESSION
AF210746
VERSION
AF210746.1 GI:6563338
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
AF210746 2160 bp DNA linear BCT 17-SEP-2001
Streptococcus pneumoniae isolate SP 1513 penicillin-binding protein
1A (pbpla) gene, complete cds.
Streptococcus pneumoniae.
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus pneumoniae.
1 (bases 1 to 2160)
Ferroni, A. and Berche, P.
Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
penicillin-resistant clinical isolates of Streptococcus pneumoniae
serotype 23F from the nasopharyngeal flora of children
J. Med. Microbiol. 50 (9), 828-832 (2001)
21432820
11549185
2 (bases 1 to 2160)
Ferroni, A. and Berche, P.
Direct Submission
Submitted (03-DEC-1999) Microbiology, Hospital
Necker-Enfants-Malades, 149 rue de Sevres, Paris 75015, France
Location/Qualifiers
1. .2160
/organism="Streptococcus pneumoniae"

		/serotype="23F"			
		/isolate="sp_1513"			
gene		/db_xref="taxon:1313"			
		1. .2160			
CDS		/gene="pbpla"			
		1. .2160			
		/gene="pbpla"			
		/codon_start=1			
		/transl_table=11			
		/product="penicillin-binding protein 1A"			
		/protein_id="AAFI7256.1"			
		/db_xref="GI:5563339"			
		/translation="MNKPTILRLIKYLSIFLSLVIAAIVLGGGVFFYYVSKAPLSLSE SKLVATSSRYNDKNLIADLSERRVNAQNDIPTDLVKAIVSIEDHRRFDRHIGID TIRILGATLRNLSQSGSTLTQOLIKLTFSTSDQTSKRAQAWLAIOLEQK ATKORILVYINKVYMSNGYMGOTAAQNYGKDLNLSLPOLALLAGCAPOAQNQYDP YSHPAQDRNRLVSEMNQGYISAEQYKAVNTPITDGLQSLKASNPYAMDNYL KEVINQVEETGNLLTGMVYINVDQEAQKHLWDIYNTEYYAYPDELOQVASTIV DVSNGKVIQAQGHQSSNVFNGINQAVETNRDWSMKPIITDVALEYGYVDSTAT IVHDEPNYPGTNPVYVNDRGYFNGINTLOALQOSRNVPAVETLNKYLNRKFTFLN GLGIDYPSIHSYNATSSNTTESDKYASSEKMAAAYAAFAANGTYIYKPMYIHKVVF DGSEKEFSNVGTRAMKETAYMMTEMMKTLSYGRNAYLAWLPQACKTGTSTNYTDE ELENHIKTSQVAPDELFAGYTRYKSYMAVWTGYSNRLPLVGNGLTVAAYKYRSMWTY LSEGNPDWNIPEELRYNGEFVFKNGARSTWNSPAPOPPSTESSSSSSSSSTSSQSS STTSTNNSTNTNPNNTTQSQNTTDPDQONPQPAQP"			
BASE COUNT		700	a	518	c
ORIGIN		410	g	532	t
Query Match 74.5%; Score 1489; DB 1; Length 2160;					
Best Local Similarity 99.5%; Pred. No. 0;					
Matches 1989; Conservative 0; Mismatches 10; Indels 0; Gaps 0;					
Qy	1	TAAATCTACGACAATAAAATCAACTCATTCGTGACTTGGTCTCTGAACGCGCGTCAA	60		
Db	159	TAAATCTACGACAATAAAATCAACTCATTCGTGACTTGGTCTCTGAACGCGCGTCAA	218		
Qy	61	TGCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA	120		
Db	219	TGCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA	278		
Qy	121	TCGCTTCTTCGACACAGGGGATATGATACCATCGGTATCTTCGGGAGCTTCTTCGCGAA	180		
Db	279	TCGCTTCTTCGACACAGGGGATATGATACCATCGGTATCTTCGGGAGCTTCTTCGCGAA	338		
Qy	181	TCGCAAAAGCAATTCCTCCAAAGTGGATCAACTCTCACCCAAACAGTTGATTAAAGTTGAC	240		
Db	339	TCGCAAAAGCAATTCCTCCAAAGTGGATCAACTCTCACTCAACAGTTGATTAAAGTTGAC	398		
Qy	241	TTACTTTTCAACTTCGACTTCCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	300		
Db	399	TTACTTTTCAACTTCGACTTCCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	458		
Qy	301	AGCGATTTCAGTTAGACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA	360		
Db	459	AGCGATTTCAGTTAGACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA	518		
Qy	361	GGTCTACATGCTAATGGGAATATGGAATGCAGACAGAGCTCAAACTACTATGGTAA	420		
Db	519	GGTCTACATGCTAATGGGAATATGGAATGCAGACAGAGCTCAAACTACTATGGTAA	578		
Qy	421	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTAGCCTGGCAATGCCTCAGGCACC	480		
Db	579	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTAGCCTGGCAATGCCTCAGGCACC	638		
Qy	481	AAACCAATATGACCCCTATTACATCCAGACAGAGCCCCAAGACCCGCCAAATCTTGGTCTT	540		
Db	639	AAACCAATATGACCCCTATTACATCCAGACAGAGCCCCAAGACCCGCCAAATCTTGGTCTT	698		
Qy	541	ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	600		
Db	699	ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	758		

Qy	601	ACCAATTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATCGA	666		
Db	759	ACCAATTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATCGA	818		
Qy	661	TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGAACAGCGCTATAACCTTACTCAC	720		
Db	819	TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGAACAGCGCTATAACCTTACTCAC	878		
Qy	721	AACTGGGATGGATGTTCTTACAAAAATGTAGACCAGAAAGAGCTCAAAAAACATCTTGTGGATAT	780		
Db	879	AACTGGGATGGATGTTCTTACAAAAATGTAGACCAGAAAGAGCTCAAAAAACATCTTGTGGATAT	938		
Qy	781	TTACAATACAGACGATACGCTTGCTATCCAGACGATGAATTGCAAGTTCGCTTCTTACCAT	840		
Db	939	TTACAATACAGACGATACGCTTGCTATCCAGACGATGAATTGCAAGTTCGCTTCTTACCAT	998		
Qy	841	TGTTTGATGTTTCTTAACGGGTAAAGTCATTGCCAGCTAGGAGCAGCCATCACTCAAGTAA	900		
Db	999	TGTTTGATGTTTCTTAACGGGTAAAGTCATTGCCAGCTAGGAGCAGCCATCACTCAAGTAA	1058		
Qy	901	TGTTTCTCTCGGAATTAACCAAGCAGTAGAAAAACAACCGGACCTGGGGATCAACATATGAA	960		
Db	1059	TGTTTCTCTCGGAATTAACCAAGCAGTAGAAAAACAACCGGACCTGGGGATCAACATATGAA	1118		
Qy	961	ACCGATCACAGACTATGCTCCTGCTTGGAGTAGCGGTGTCTACGATTCAACTGCTACTAT	1020		
Db	1119	ACCGATCACAGACTATGCTCCTGCTTGGAGTAGCGGTGTCTACGATTCAACTGCTACTAT	1178		
Qy	1021	CGTTCCAGATGAGCCCTATAACTACCCTGGGACAAATATCTCTGTTTATACTGGGATAG	1080		
Db	1179	CGTTCCAGATGAGCCCTATAACTACCCTGGGACAAATATACCCCTGTTTATACTGGGATAG	1238		
Qy	1081	GGCTTACTTTGGCAACATCACTTGCATACGCCCTCGCAAAATCGGAAACGTCGCCAGC	1140		
Db	1239	GGCTTACTTTGGCAACATCACTTGCATACGCCCTCGCAAAATCGGAAACGTCGCCAGC	1298		
Qy	1141	CGTGGAACTCTAAACAAGTCGGACTCAACCGCGCCCAAGACTTTCCTAAATGCTCTAGG	1200		
Db	1299	CGTGGAACTCTAAACAAGTCGGACTCAACCGCGCCCAAGACTTTCCTAAATGCTCTCGG	1358		
Qy	1201	AATCGACTACCCAAAGTATTCACTACTCAATGCCATTTCAGTAAACACACCGAATCAGA	1260		
Db	1359	AATCGACTACCCAAAGTATTCACTACTCAATGCCATTTCAGTAAACACACCGAATCAGA	1418		
Qy	1261	CAAAAAATATGGAGCAAGTAGTGAAGAGATGGCTGCTCTTACGCTGCTTTTGCAAAATGG	1320		
Db	1419	CAAAAAATATGGAGCAAGTAGTGAAGAGATGGCTGCTCTTACGCTGCTTTTGCAAAATGG	1478		
Qy	1321	TGGAACCTTACTATAAACCAATGTATATCCATAAAGTCGTCTTTTAGTATGGGAGTGA	1380		
Db	1479	TGGAACCTTACTATAAACCAATGTATATCCATAAAGTCGTCTTTTAGTATGGGAGTGA	1538		
Qy	1381	AGAGTTCTCTAATGTCGGAACCTCGTGCCATGAGGAAACGACACGCTTATATGATGACCGA	1440		
Db	1539	AGAGTTCTCTAATGTCGGAACCTCGTGCCATGAGGAAACGACACGCTTATATGATGACCGA	1598		
Qy	1441	CATGATGAAACACAGTCTTGACTTATGGAACCTGGACGAAATGCCTTATCTTGCTTGGCTCC	1500		
Db	1599	AATGATGAAACACAGTCTTGAGTTATGGAACCTGGACGAAATGCCTTATCTTGCTTGGCTCC	1658		
Qy	1501	TCAGGCTGGTAAACAGGAACCTCTTAACATACAGACGAGGAAATTTGAAAACCATATCAA	1560		
Db	1659	TCAGGCTGGTAAACAGGAACCTCTTAACATACAGATGAGGAAATTTGAAAACCATATCAA	1718		
Qy	1561	GACCTCTCAATTTGTAGCACCTGATGAACATTTTGTGCGGTATACGCGTAAATATTTCAAT	1620		
Db	1719	GACCTCTCAATTTGTAGCACCTGATGAACATTTTGTGCGGTATACGCGTAAATATTTCAAT	1778		
Qy	1621	GGCTGTATGACAGGCTATTCTTAACGCTCTGACACCACTTCTAGCGAATGCGCTTACGCT	1680		
Db	1779	GGCTGTATGACAGGCTATTCTTAACGCTCTGACACCACTTCTAGCGAATGCGCTTACGCT	1838		
Qy	1681	CGCTGCAAAAGTTTACCGCTCTATGATGAGCTACCTGCTGTAAGGAAGCAATCCAGAAGA	1740		

Qy	601	AC	CAATTA	CTGATGGACTACAAAGTCTCAAAATAGCAAGTAATTAACCTGCTTACATGGA	660
Db	759	AC	CAATTA	CTGATGGACTACAAAGTCTCAAAATAGCAAGTAATTAACCTGCTTACATGGA	818
Qy	661	TA	ATTACCT	CAAGGAAGTCAATCAATCAAGTTGAAGAGAAACAGGCTATAACCTACTCAC	720
Db	819	TA	ATTACCT	CAAGGAAGTCAATCAATCAAGTTGAAGAGAAACAGGCTATAACCTACTCAC	878
Qy	721	AA	CTGGGAT	GGATGCTCACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780
Db	879	AA	CTGGGAT	GGATGCTCACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	938
Qy	781	TT	ACAATAC	CAGACGAATAGTTGCTTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT	840
Db	939	TT	ACAATAC	CAGACGAATAGTTGCTTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT	998
Qy	841	TG	TTGATG	TCTTAACGGTAAAGTCAATGCTCCAGCTAGGAGCAGCGCATCAGTCAAGTAA	900
Db	999	TG	TTGATG	TCTTAACGGTAAAGTCAATGCTCCAGCTAGGAGCAGCGCATCAGTCAAGTAA	1058
Qy	901	TG	TTTCTCT	TCGGAATTAACCAAGCAGTACAAACAAACCGGACTGGGGATCAACTATGAA	960
Db	1059	TG	TTTCTCT	TCGGAATTAACCAAGCAGTACAAACAAACCGGACTGGGGATCAACTATGAA	1118
Qy	961	AC	CGATAC	CAGACTATGCTCTGCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT	1020
Db	1119	AC	CGATAC	CAGACTATGCTCTGCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT	1178
Qy	1021	CG	TTACGAT	GAGCCCTATAACTACCTGGGCAATACTCTCTGTTTATAACTGGGATAG	1080
Db	1179	CG	TTACGAT	GAGCCCTATAACTACCTGGGCAATACTCTCTGTTTATAACTGGGATAG	1238
Qy	1081	GG	GTACTT	TGGCAACATCACCTTGCATACGCCCTGCAACCAATCGGAAACGCTCCAGC	1140
Db	1239	GG	GTACTT	TGGCAACATCACCTTGCATACGCCCTGCAACCAATCGGAAACGCTCCAGC	1298
Qy	1141	CG	TGAAACT	CTAAACAAGTGGAGTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG	1200
Db	1299	CG	TGAAACT	CTAAACAAGTGGAGTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG	1358
Qy	1201	AA	TCGACT	TACCCCAAGTATTCACCTACTCAATGCCATTTCAAGTAAACACACCAATCAGA	1260
Db	1359	AA	TCGACT	TACCCCAAGTATTCACCTACTCAATGCCATTTCAAGTAAACACACCAATCAGA	1418
Qy	1261	CA	AAAAAT	TGAGCAAGTAGTGAAAGATGGCTGCTGTCTAGCTGCTCTTTCGCAATGG	1320
Db	1419	CA	AAAAAT	TGAGCAAGTAGTGAAAGATGGCTGCTGTCTAGCTGCTCTTTCGCAATGG	1478
Qy	1321	TG	GAATCT	ACTATAAAACCAATGTATATCCATAAAGTCGCTTTTGTAGTGGGAGTGA	1380
Db	1479	TG	GAATCT	ACTATAAAACCAATGTATATCCATAAAGTCGCTTTTGTAGTGGGAGTGA	1538
Qy	1381	AG	AGTTCT	TAATGTCGGAACCTGCTCCATGAAGGAAACGACAGCCATATGATGACCGA	1440
Db	1539	AG	AGTTCT	TAATGTCGGAACCTGCTCCATGAAGGAAACGACAGCCATATGATGACCGA	1598
Qy	1441	CA	TGATGA	AAACACAGCTTCTGACTTATGGAATGGACCAATGCTATCTTGTGGCTGCC	1500
Db	1599	AA	TGATGA	AAACACAGCTTCTGACTTATGGAATGGAACTGGACCAATGCTATCTTGTGGCTGCC	1558
Qy	1501	TC	AGGCTG	TAATAAACAGAACCTTAACATATACAGACGAGGAAATTAATAACACATCAA	1560
Db	1659	TC	AGGCTG	TAATAAACAGAACCTTAACATATACAGATGAGGAAATTAATAACACATCAA	1718
Qy	1561	GAC	CTCTCA	ATTGTTAGACACCTGATGAACATTTTGTGCTGCTATACCGGTAATAATTCAT	1620
Db	1719	GAC	CTCTCA	ATTGTTAGACACCTGATGAACATTTTGTGCTGCTATACCGGTAATAATTCAT	1778
Qy	1621	GG	CTGTAT	GCACAGGCTATTCTAACCCGCTGCACACCATTTGTAGGCAATGGCCTTACGGT	1680
Db	1779	GG	CTGTAT	GCACAGGCTATTCTAACCCGCTGCACACCATTTGTAGGCAATGGCCTTACGGT	1838
Qy	1681	CG	CTGCA	AAAGTTTACCCTCTATGATGACCTACCTGTCTGTAAGGAAGCAATCCGAAGA	1740


```
|||||
Db 1839 CCGTCGCAAGTTACCGCTCTATGATGACCTATCTGTCTGAAGGAAGCAATCCAGAGGA 1898
QY 1741 TTGGAATATACAGAGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 1800
Db 1899 TTGGAATATACAGAGGACTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 1958
QY 1801 TTCTAGGTGGAAGTACCTACCTGCTCCACACAAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
Db 1959 TTCTAGGTGGAAGTACCTACCTGCTCCACACAAACCCCATCAACTGAAAGTTCAAGCTCATC 2018
QY 1861 ATCAGATAGTTCAACTTTCACAGTCTAGCTCAACCACTCCAGCACAATAAATAGTAGCAG 1920
Db 2019 ATCAGATAGTTCAACTTTCACAGTCTAGCTCAACCACTCCAGCACAATAAATAGTAGCAG 2078
QY 1921 TACCATCTCAACAATAATACGCAACAATCAATAACAACCCCTGATCAACAAAATCAGAA 1980
Db 2079 TACCATCTCAACAATAATACGCAACAATCAATAACAACCCCTGATCAACAAAATCAGAA 2138
QY 1981 TCCTCAACCAAGCAACCA 1999
Db 2139 TCCTCAACCAAGCAACCA 2157

RESULT 7
SPON7
LOCUS SPON7 2157 bp DNA linear BCT 11-MAY-1995
DEFINITION S.pneumoniae (63915) ponA gene for penicillin-binding protein 1a.
ACCESSION X67872
VERSION X67872.1 GI:47417.
KEYWORDS penicillin-binding protein 1a; ponA gene.
SOURCE Streptococcus pneumoniae.
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.
REFERENCE 1 (bases 1 to 2157)
AUTHORS Hakenbeck, R.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek. Genetik, Ihnestr. 73, 1000 Berlin 33, FRG
REFERENCE 2 (bases 1 to 2157)
AUTHORS Martin, C., Sibold, C. and Hakenbeck, R.
TITLE Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain
JOURNAL EMBO J. 11 (11), 3831-3836 (1992)
MEDLINE 93010977
FEATURES
source
1..2157 Location/Qualifiers
/organism="Streptococcus pneumoniae"
/strain="63915"
/db_xref="taxon:1313"
/clone="pcmk1-pcmk16"
1..2157
/gene="ponA"
1..>2157
/gene="ponA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="CAA48072.1"
/db_xref="GI:47418"
/swiss="SWISS-PROT:Q04707"
/translation="MNKPTILRLIKYLSFSLVIAIILVGGGVFFVYVSKAPLSLSE
SKLVATSSKIVDNKNOLIADLSERRVNAOANDIPTDLVRAIVSIEDHRFPDHRGID
TIRLGLFRLNQSLOGSLTQOLIKLTFESTSDQISRAQEAWLAIQLEQK
ATKQELIYYINKVLSNGNTGMQTAQNYVKGDLNLSLPOLALLAGMPQNPYDP
YSHPEAARDRLVSEMKNOGYISAEQYERAVNPTITDGLQSLKSASNPAYMDNYL
KEVINQVEETGYNLLTGMVYTNVDQEAQHLWDIYNTDEYVAYPDDELOQVSTIV
DVSNGKVAQLGARHQSNSVFGINQAVETNRDWSMTKPTIDYAPALEYGYDSTAV
IVHDKPYNPTNTPVYNDRGYFGNITLOYALQOOSRVNPAVETLNKVLNPKATFLN
GLGIDYPSIHVSNALSSNTTSDKKYIGASSEKMAAAYAFANGGTYIYKPMYIHKVYFS
DGSEKFSNVGTRAIKETATYMTMDMKTVLSYGTGRNAYLAWLPQAGKTGTSNTIDE
```

```
ETENHIKTSQFVAPDELFAGYTRKYSMAVVTGYSNRLPLVNGNLUTVAAKYVRSMYTY
LSEGNPNEDWNIPEGLYRNGEEFKNGSTWNPAPQPPSTSESSSSDSTSQSS
STTPSTNNSTTTNNNTTQOSNTTDDQONONPPAPQ"
BASE COUNT 699 a 521 c 407 g 530 t
ORIGIN
Query Match 71.9%; Score 1438; DB 1; Length 2157;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 TAAATCTACGACAAATAAATAACAACCTCATCTGCTGACTTGGGTCTCAAGCCGCGTCAA 60
Db 159 TAAATCTACGACAAATAAATAACAACCTCATCTGCTGACTTGGGTCTCAAGCCGCGTCAA 218
QY 61 TGCCCAAGCTAATGATATTTCCCAACAGATTGGTTAAGGCAATCGTTCTATCAAGAACCA 120
Db 219 TGCCCAAGCTAATGATATTTCCCAACAGATTGGTTAAGGCAATCGTTCTATCAAGAACCA 278
QY 121 TCGCTTCTTCGACCAACAGGGGATTCATCCATCCGTATCTCTGGAGCTTTCTTTCGCGAA 180
Db 279 TCGCTTCTTCGACCAACAGGGGATTCATCCATCCGTATCTCTGGAGCTTTCTTTCGCGAA 338
QY 181 TCTGCAAGCAATTCCTCCAAAGTGGATCACTCTCACCACACAGTTGATTAAAGTTGAC 240
Db 339 TCTGCAAGCAATTCCTCCAAAGTGGATCACTCTCACCACAGTTGATTAAAGTTGAC 398
QY 241 TTACTTTTCAACTTCGACTTCCGACCAAGACTATTTCTCGTAAGGCTCAGGAAGCTGGTT 300
Db 399 TTACTTTTCAACTTCGACTTCCGACCAAGACTATTTCTCGTAAGGCTCAGGAAGCTGGTT 458
QY 301 AGGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTCACCTACTATATAATAA 360
Db 459 AGGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTCACCTACTATATAATAA 518
QY 361 GGTCTACATCTCTAATGGGAAGTATGGAATGCAGACAGAGCTCAAACTACTATATGTA 420
Db 519 GGTCTACATCTCTAATGGGAAGTATGGAATGCAGACAGAGCTCAAACTACTATATGTA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGATGCTTGGCTCAGACACC 480
Db 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGATGCTTGGCTCAGACACC 638
QY 481 AAACCAATATGACCCCTATTCATCCAGAGAGCCCAAGAGCCGCGGAAACTTGGTCTT 540
Db 639 AAACCAATATGACCCCTATTCATCCAGAGAGCCCAAGAGCCGCGGAAACTTGGTCTT 698
QY 541 ATCTGAATCAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTCAATAC 600
Db 699 ATCTGAATCAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTCAATAC 758
QY 601 ACCAATCTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGGA 660
Db 759 ACCAATCTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAAGCAAGGCTATTAACCTACTCAC 720
Db 819 TAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAAGCAAGGCTATTAACCTACTCAC 878
QY 721 AACTGGGATGGATGCTACACAAATCTAGACCAAGAGCTCAAAACATCTGTGGGATAT 780
Db 879 AACTGGGATGGATGCTACACAAATCTAGACCAAGAGCTCAAAACATCTGTGGGATAT 938
QY 781 TTACAATACAGCAATAACCTTGCCTATCCAGACGATGAATTCAGAGTCGCTTCTACCAT 840
Db 939 TTACAATACAGCAATAACCTTGCCTATCCAGACGATGAATTCAGAGTCGCTTCTACCAT 998
QY 841 TGTGTAGTTTCAACGGTAAAGTCAATGCCCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
Db 999 TGTGTAGTTTCAACGGTAAAGTCAATGCCCCAGCTAGGAGCAGCCATCAGTCAAGTAA 1058
QY 901 TGTTCCTTCGGAAATTAACCAACAGCAGTAGAACAACCCGAGCTGGGGATCAACTATGA 960
```


Db	1059	TGTTTCCTTCGGAAATTAAACAAGCAGTAGAACAACAACCGGAGCTGGGGATCAACTATGAA	1111
Qy	961	ACCGATCACAGACTATGCTCCTCGCTTGAGTAGCGGTGCTACGATTCAACTGCTACTAT	1020
Db	1119	ACCGATCACAGACTATGCTCCTCGCTTGAGTAGCGGTATCTACGATTCAACTGCTACTAT	1178
Qy	1021	CGTTCAGGATGAGCCCTATAACTACCTCCCTGGGACAANAATCTCCTGTTTATTAACCTGGGATAG	1080
Db	1179	CGTTTCAGATAAGCCCTATAAATACCTCCCTGGGACAANAATACCCCTGTTTATAACTGGGATAG	1238
Qy	1081	GGGCTACTTTTGGCAACAATCAACCTTGCAATPACGCCCTGCAACAATCGGAAACGTCGCCAGC	1140
Db	1239	GGGCTACTTTTGGCAACAATCAACCTTGCAATPACGCCCTGCAACAATCGGAAACGTCGCCAGC	1298
Qy	1141	CGTGGAAACTCTTAACAAGGTGGGACTCAACCGCGCCCAAGACTTTTCCTTAATATGTCCTAGG	1200
Db	1299	CGTGGAAACTCTTAACAAGGTGGGACTCAACCGCGCCCAAGACTTTTCCTTAATATGTCCTCGG	1358
Qy	1201	AATCGACTACCCCAAGTATTCACTACTCAAAATGCCATTTCAAAGTAAACAACCGGAATCAGA	1260
Db	1359	AATCGACTACCCCAAGTATTCACTACTCAAAATGCCATTTCAAAGTAAACAACCGGAATCAGA	1418
Qy	1261	CAAAAAATATGGAGCAAGTAGTGAAGAATGCGTCTGCTTACGCTGCTTTTGCAAAATGG	1320
Db	1419	CAAAAAATATGGAGCAAGTAGTGAAGAATGCGTCTGCTTACGCTGCTTTTGCAAAATGG	1478
Qy	1321	TGGAACTTACTATAAACCAATGTATATCCATAAAGTCTCTTTAGTGNATGGGAGTGAAAA	1380
Db	1479	TGGAACTTACTATAAACCAATGTATATCCATAAAGTCTCTTTAGTGNATGGGAGTGAAAA	1538
Qy	1381	AGAGTTCTCTTAATGTTCGGAACTCGTGCCATGAAGAAACGACAGCCCTATATGATGACCCGA	1440
Db	1539	AGAGTTCTCTTAATGTTCGGAACTCGTGCCATGAAGAAACGACAGCCCTATATGATGACCCGA	1598
Qy	1441	CATGATGAAAAAGTCTTGACTTATGGAACTCGGACGAATGCCTATCTTGCTTGGGCTCCC	1500
Db	1599	CATGATGAAAAAGTCTTGACTTATGGAACTCGGACGAATGCCTATCTTGCTTGGGCTCCC	1658
Qy	1501	TCAGGCTGGTAAAAACAGGAACCTCTAACTATATACAGACGAGGAATTTGAAAAACCAATCAA	1560
Db	1659	TCAGGCTGGTAAAAACAGGAACCTCTAACTATATACAGACGAGGAATTTGAAAAACCAATCAA	1718
Qy	1561	GACCTCTCAATTTGTAGCAGCTGATGAATCTTTGCTGGCTATACGCGTAATAATATTCAT	1620
Db	1719	GACCTCTCAATTTGTAGCAGCTGATGAATCTTTGCTGGCTATACGCGTAATAATATTCAT	1778
Qy	1621	GGCTGTATGACAGGCTATTCTTAACCGTCTGACACCACTTTGTAGGCAATGSCCTTACGGT	1680
Db	1779	GGCTGTATGACAGGCTATTCTTAACCGTCTGACACCACTTTGTAGGCAATGSCCTTACGGT	1838
Qy	1681	CGCTGCCAAAGTTTACCCTCTATGATGACCTTACCTGTCTGAAGGAAGCAATCCAGAGA	1740
Db	1839	CGCTGCCAAAGTTTACCCTCTATGATGACCTTACCTGTCTGAAGGAAGCAATCCAGAGA	1898
Qy	1741	TTGGAATATACCAGAGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTGTCTCG	1800
Db	1899	TTGGAATATACCAGAGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTGTCTCG	1958
Qy	1801	TTCTACGTGAACCTACCTCGCTCCACAACACCCCCATCAACTGAAAGTTTCAAGCTCATC	1860
Db	1959	TTCTACGTGAACCTACCTCGCTCCACAACACCCCCATCAACTGAAAGTTTCAAGCTCATC	2018
Qy	1861	ATCAGATAGTTTCAACTTTCAGCTTACGCTACACCACTCCAAAGCACAATATATAGTAGCAC	1920
Db	2019	ATCAGATAGTTTCAACTTTCAGCTTACGCTACACCACTCCAAAGCACAATATATAGTAGCAC	2078
Qy	1921	TACCAATCTTAAACAATTAATAGCAACAATCAAAATACAACCCCTGATCAACAATAATCAGAA	1980
Db	2079	TACCAATCTTAAACAATTAATAGCAACAATCAAAATACAACCCCTGATCAACAATAATCAGAA	2138
Qy	1981	TCCTCAACCAAGCACAACCA	1999
Db	2139	TCCTCAACCAAGCACAACCA	2157


```
QY 601 ACCAATTACTGATGAGTACTCAAAAGTCTCAAAATCAGCAAGTAATATACCTGCTTACATGGA 660
Db 759 ACCAATTACTGATGGGCTACAAAGTCTCAAAATCAGCAAGTAATATACCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAGAACAAGGCTATTAACCTACTAC 720
Db 819 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAGAACAAGGCTATTAACCTACTAC 878
QY 721 AACTGGGATGGATGTTCTACAAATATAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
Db 879 AACTGGGATGGATGTTCTACAAATATAGACCAAGAGCTCAAAAACATCTGTGGGATAT 938
QY 781 TTACAATACAGAGCAATACCTGGCTATCCAGACGATGAATGCAAGTGCCTTCTACCAT 840
Db 939 TTACAATACAGAGCAATACCTGGCTATCCAGACGATGAATGCAAGTGCCTTCTACCAT 998
QY 841 TGTGTGATGTTCTTAACGGTAAAGTCATTTGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
Db 999 TGTGTGATGTTCTTAACGGTAAAGTCATTTGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 1058
QY 901 TGTTCCTTCGGAATTAAACCAAGCAGTAGAAACAAACCCGAGCTGGGGATCAACTATGAA 960
Db 1059 TGTTCCTTCGGAATTAAACCAAGCAGTAGAAACAAACCCGAGCTGGGGATCAACTATGAA 1118
QY 961 ACCGATCACAGACTATGCTCCTGCTTGGAGTAGCGGTGCTCTACGATTCACCTGCTACTAT 1020
Db 1119 ACCGATCACAGACTATGCTCCTGCTTGGAGTAGCGGTGCTCTACGAGTCAACTGCGCACT 1178
QY 1021 CGTTCAGGATGAGCCCTATAACTACCTCGGACAAATACTCCTGTTTATAACTGGGATAG 1080
Db 1179 CGTTCAGGATGAGCCCTATAACTACCTCGGACAAATACTCCTGTTTATAACTGGGATAG 1238
QY 1081 GGCGTACTTTGGCAACATCACTTGCATACGCGCTCGCAATCGGGAACCGTCCCGAGC 1140
Db 1239 GGCGTACTTTGGCAACATCACTTGCATACGCGCTCGCAATCGGGAACCGTCCCGAGC 1298
QY 1141 CGTGGAACCTCTAAACAGTCGAGCTCAACCGCGCAAGACTTTCTAAATGGTCTAGG 1200
Db 1299 CGTGGAACCTCTAAACAGTCGAGCTCAACCGCGCAAGACTTTCTAAATGGTCTCGG 1358
QY 1201 AATCGACTACCCCAAGTATTCACACTACTCAATGCCATTCAGTAACCAACCGCAATCAGA 1260
Db 1359 AATCGACTACCCCAAGTATTCACACTACTCAATGCCATTCAGTAACCAACCGCAATCAGA 1418
QY 1261 CAAAAATATGGAGCAAGTAGTGAAGAAGATGGCTGCTGTCTAGCTGCCTTTTGCAAATGG 1320
Db 1419 CAAAAATATGGAGCAAGTAGTGAAGAAGATGGCTGCTGTCTAGCTGCCTTTTGCAAATGG 1478
QY 1321 TGGAACTTACTATAAACCAATGATATCCATAAAGTCTCTTTAGTGATGGGAGTGAATA 1380
Db 1479 TGGAACTTACTATAAACCAATGATATCCATAAAGTCTCTTTAGTGATGGGAGTGAATA 1538
QY 1381 AGAGTCTCTTAATGTCGGAACCTCGTGCCATGAAGGAACACAGCCCTATATGATGACCGA 1440
Db 1539 AGAGTCTCTTAATGTCGGAACCTCGTGCCATGAAGGAACACAGCCCTATATGATGACCGA 1598
QY 1441 CATGATGAAACAGTCTTGACTATGGAACCTGGACGAATAGCCCTATCTTCTTGGCTCCC 1500
Db 1599 CATGATGAAACAGTCTTGACTATGGAACCTGGACGAATAGCCCTATCTTCTTGGCTCCC 1658
QY 1501 TCAGGCTGGTAAACAGGAACCTCTACTATACAGAGGAGGAATTGAAACACACATCAA 1560
Db 1659 TCAGGCTGGTAAACAGGAACCTCTACTATACAGAGGAGGAATTGAAACACACATCAA 1718
QY 1561 GACCTCTCAATTTGTAGCACCTGATGACATATTTGCTGGCTATACGCGTAAATATTCAT 1620
Db 1719 GACCTCTCAATTTGTAGCACCTGATGACATATTTGCTGGCTATACGCGTAAATATTCAT 1778
QY 1621 GCGTGTATGACAGGCTTATCTAACCGTCTGACACCACTGTGAGGCAATGSCCTTACGGT 1680
Db 1779 GCGTGTATGACAGGCTTATCTAACCGTCTGACACCACTGTGAGGCAATGSCCTTACGGT 1838
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTGTGAAGGAAGCAATCCAGAAGA 1740
```

```
Db 1839 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTGAAGGAAGCAATCCAGAGGA 1898
QY 1741 TTGGAATATACCAGAGGGGCTCTACAGAAATGGAGAATTCGTATTTAAAAATGCTGCTCG 1800
Db 1899 TTGGAATATACCAGAGGGGCTCTACAGAAATGGAGAATTCGTATTTAAAAATGCTGCTCG 1958
QY 1801 TTCTACGTGGAACCTACGCTGCTCCACAAACACCCCCATCAACTGAAAGTTTCAAGCTCATC 1860
Db 1959 TTCTACGTGGAACCTACGCTGCTCCACAAACACCCCCATCAACTGAAAGTTTCAAGCTCATC 2018
QY 1861 ATCAGATAGTTCAACTTCAAGTCTAGCTCAACCACTCCCAAGCACAATAATAGTAGAC 1920
Db 2019 ATCAGATAGTTCAACTTCAAGTCTAGCTCAACCACTCCCAAGCACAATAATAGTAGAC 2078
QY 1921 TACCAATCCTTAACAATAATACGCAACATCAATACCAACCCCTGATCAACAAAAATCAGAA 1980
Db 2079 TACCAATCCTTAACAATAATACGCAACATCAATACCAACCCCTGATCAACAAAAATCAGAA 2138
QY 1981 TCCTCAACACGACACAACCA 1999
Db 2139 TCCTCAACACGACACAACCA 2157

RESULT 9
STRPONA
LOCUS
DEFINITION STRPONA 3378 bp DNA linear BCT 26-APR-1993
Streptococcus pneumoniae penicillin-binding protein (ponA) gene,
complete cds.
ACCESSION M90527.1 GI:153766
VERSION M90527.1 GI:153766
KEYWORDS penicillin-binding protein; ponA gene.
SOURCE Streptococcus pneumoniae (strain R6) DNA.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 3378)
AUTHORS Martin,C., Briese,T. and Hakenbeck,R.
TITLE Nucleotide sequences of genes encoding penicillin-binding proteins
from Streptococcus pneumoniae and Streptococcus oralis with high
homology to Escherichia coli penicillin-binding proteins 1A and 1B
J. Bacteriol. 174, 4517-4523 (1992)
JOURNAL
MEDLINE
FEATURES
source Location/Qualifiers
1..3378
/organism="Streptococcus pneumoniae"
/strain="R6"
/db_xref="taxon:1313"
315..3105
/gene="ponA"
315..320
/gene="ponA"
353..949
/gene="ponA"
/function="unknown"
/notes="ORF"
/codon_start=1
/transl_table=11
/protein_id="AAA26955.1"
/db_xref="GI:153767"
/translation="MVNYPHKVSSQKROTSLSQKFNFRGMSFEKMINATNDYILSQ
GLAIVHKPTQIOIVQVDYQPSRAKIVFQASQTDYSGVYNGYIDFVKETKQ
KRAIPKMFPHQIOHMEQVLAQGGICFVLLHPFSQOETVLLPAFDLIRFYHODKQK
SMPLVIREYGYEIKAGAFPQIPYLVNVIKHLGGKTR"
946..3105
/gene="ponA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein"
/protein_id="AAA26956.1"
/db_xref="GI:153768"
/translation="MNKPTILRLIKYLSIFLSLVIAAIVLGGGVFFYVSKAPSLSE
SKLVATSSSKIYDNKNQLIADLGSRRVNAQANDIPTDLVKAIVSIEDHRFDFHRGID
```


TIRILGAFURNLOSLSQGSALTQOLIKLTVSTSTSDTISRKAQEWLAIOLEQK
ATKQELTYIIINKVYMSNGNYGKDLNNLSLPOLALLAGMPQPNYDP
YGRPAADORRNVLISEMKNQYISAEQYKAVNPITDGLQSLKASASNYPAYMDNYL
KEVINQVEETCYNLLTTCMDVYTNVDOEAQKHLMDIYNDEVVAYPDELOVASTIV
DYSNGKVIQAQLGARQSSNVSEFINQAVETNRDWSGTMKPIIDYAPALEYGYVESTAT
IVHDEPINTPGNTPIYNWDRGYFNITLOIALQSSRNPAVETLINKVGLNRAKTFLN
GLGIDPYSNYSNAISSNTTESDKKYGASSEKMAAAYAFANGGYIYKPMYTHKVYFS
DGESEKFSNVGRAMKETAYNMKTMLVLSYGTGRNAYLAWLPOAGKTYSNYTD
ETENHKTISOFAPELFAGYTRKYSMVMTGYSNRLTPLVNGLITVAKVYRSMWTY
LSEGSNPEDMNIPEGLYRNGEFVKNGASTWSSPAQOPPSTESSSSSDSSTSOSS
STTPSNNSNTTNPNNNTQOOSNTTDDOQNQNPQAP"

BASE COUNT	1093 a	783 c	609 g	893 t	
ORIGIN					
Query Match	71.9%;	Score 1438;	DB 1;	Length 3378;	
Best Local Similarity	99.4%;	Pred. No. 0;			
Matches 1988;	Conservative 0;	Mismatches 11;	Indels 0;	Gaps 0;	
QY 1	TAAATCTAGCACAATAAATAATCAATCTGCTGACTTGGGTCTTGAACGCCGCGTCAA	60			
DB 1104	TAAATCTAGCACAATAAATAATCAATCTGCTGACTTGGGTCTTGAACGCCGCGTCAA	1163			
QY 61	TGCCCAAGCTAATGATATTCGCCACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA	120			
DB 1164	TGCCCAAGCTAATGATATTCGCCACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA	1223			
QY 121	TCGCTTCTTCGACACAGGGGGATTGATACCATCCGTATCTCGGAGCTTCTTTCGGCAA	180			
DB 1224	TCGCTTCTTCGACACAGGGGGATTGATACCATCCGTATCTCGGAGCTTCTTTCGGCAA	1283			
QY 181	TCGCAAGCAATTCCTCCAAAGTGGATGATCACTCAACCAAGTTGATTAAAGTTGAC	240			
DB 1284	TCGCAAGCAATTCCTCCAAAGTGGATGATCACTCAACCAAGTTGATTAAAGTTGAC	1343			
QY 241	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAACCTTGGTT	300			
DB 1344	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAACCTTGGTT	1403			
QY 301	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTTACTATATAATAA	360			
DB 1404	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTTACTATATAATAA	1463			
QY 361	GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAAACTACTATGGTAA	420			
DB 1464	GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAAACTACTATGGTAA	1523			
QY 421	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGATGCCTCAGGCACC	480			
DB 1524	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGATGCCTCAGGCACC	1583			
QY 481	AAACCAATATGACCCCTATTTCACATCCAGAAAGCAGCCCAAGACCGCGGAAACTTGGTCTT	540			
DB 1584	AAACCAATATGACCCCTATTTCACATCCAGAAAGCAGCCCAAGACCGCGGAAACTTGGTCTT	1643			
QY 541	ATCTGAATGAAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	600			
DB 1644	ATCTGAATGAAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	1703			
QY 601	ACCAATTTACTGTGGACTACAAAGTCTCAATCAGCAAGTAATTACCTGCTTACATGGA	660			
DB 1704	ACCAATTTACTGTGGACTACAAAGTCTCAATCAGCAAGTAATTACCTGCTTACATGGA	1763			
QY 661	TAATTAACCTCAAGGAAGTTCATCAATCAAGTTGAAGAAAGCAAGGCTATTAACCTACTCAC	720			
DB 1764	TAATTAACCTCAAGGAAGTTCATCAATCAAGTTGAAGAAAGCAAGGCTATTAACCTACTCAC	1823			
QY 721	AACTGGGATGGATGCTTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780			
DB 1824	AACTGGGATGGATGCTTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	1883			
QY 781	TTACAATACAGCAATAGTTCGCTTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT	840			

DB 1884	TTACAAATACAGACGAATACGTTGCCTATCCAGACGATGAATTCGAAGTCGCTTCTACCAT	1943
QY 841	TGTTGATGTTTCTAAACGGTAAAGTCATTGCCAGCTAGGACGACGCCATCAGTCAAGTAA	900
DB 1944	TGTTGATGTTTCTAAACGGTAAAGTCATTGCCAGCTAGGACGACGCCATCAGTCAAGTAA	2003
QY 901	TGTTTCTCTCGGAATTAACCAAGCAGTAGAAAAACAAACCGGACTGGGGATCAACTATGAA	960
DB 2004	TGTTTCTCTCGGAATTAACCAAGCAGTAGAAAAACAAACCGGACTGGGGATCAACTATGAA	2063
QY 961	ACCGATACAGACTATGCTCCTCGCTTGGAGTAGGGTGTCTACGATTCACATGCTACTAT	1020
DB 2064	ACCGATACAGACTATGCTCCTCGCTTGGAGTAGGGTGTCTACGATTCACATGCTACTAT	2123
QY 1021	CGTTACAGGATGAGCCCTATAACTACCTGGGACAAATACCTCTTTTATTAACATGGGATAG	1080
DB 2124	CGTTACAGGATGAGCCCTATAACTACCTGGGACAAATACCCCTGTTTATTAACATGGGATAG	2183
QY 1081	GGCTACTTTTGGCAACATCACTTTCGAATACGCCCTGCAACAATTCGGGAACCGTCCGAGC	1140
DB 2184	GGCTACTTTTGGCAACATCACTTTCGAATACGCCCTGCAACAATTCGGGAACCGTCCGAGC	2243
QY 1141	CGTGGAACTCTTAAACAAGTCGAGCTCAACCGGCCCAAGACTTTCTTAATGGTCTAGG	1200
DB 2244	CGTGGAACTCTTAAACAAGTCGAGCTCAACCGGCCCAAGACTTTCTTAATGGTCTCGG	2303
QY 1201	AATCGACTACCCCAAGTATTCACTACTCAATGCAATTCGAATTAAGTAACAACCAACCAATCAGA	1260
DB 2304	AATCGACTACCCCAAGTATTCACTACTCAATGCAATTCGAATTAAGTAACAACCAACCAATCAGA	2363
QY 1261	CAAAAAATATGGAGCAAGTAGTGAAAGATGGCTGTGCTTTAGCTGCTCTTTGCAAAATGG	1320
DB 2364	CAAAAAATATGGAGCAAGTAGTGAAAGATGGCTGTGCTTTAGCTGCTCTTTGCAAAATGG	2423
QY 1321	TGGAATCTACTATAAACCAATGATATCCATAAAGTCGCTTTTAGTCATCGGAGTGAATA	1380
DB 2424	TGGAATCTACTATAAACCAATGATATCCATAAAGTCGCTTTTAGTCATCGGAGTGAATA	2483
QY 1381	AGAGTTCTCTTAATGTGCGAACTCGTCCATGAAGAAACACAGACCTTATATGATGACCGA	1440
DB 2484	AGAGTTCTCTTAATGTGCGAACTCGTCCATGAAGAAACACAGACCTTATATGATGACCGA	2543
QY 1441	CATGATGAAACAGTCTTTGACTTATGGAAGTGCAGCAAAATGCCTATCTTCTGCTGGCTCCC	1500
DB 2544	CATGATGAAACAGTCTTTGAGTTATGGAAGTGCAGCAAAATGCCTATCTTCTGCTGGCTCCC	2603
QY 1501	TCAGGCTGGTAAACAGGAACTCTTAACATACAGACGAGGAAATTCGAAACCAACATCAA	1560
DB 2604	TCAGGCTGGTAAACAGGAACTCTTAACATACAGACGAGGAAATTCGAAACCAACATCAA	2663
QY 1561	GACCTCTCAATTTGTAGCACCTGATGAACATTTTGTGCTGCTATACCGGTAATATTCAAT	1620
DB 2664	GACCTCTCAATTTGTAGCACCTGATGAACATTTTGTGCTGCTATACCGGTAATATTCAAT	2723
QY 1621	GGCTGATGGACAGGCTATTCTAACCGCTGACACCACTTGTAGGCAATGGCCTTACGGT	1680
DB 2724	GGCTGATGGACAGGCTATTCTAACCGCTGACACCACTTGTAGGCAATGGCCTTACGGT	2783
QY 1681	CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTCTGAAGGAGCAATCCAGAGA	1740
DB 2784	CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTCTGAAGGAGCAATCCAGAGA	2843
QY 1741	TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGATTTTAAAAATGGTGTCTG	1800
DB 2844	TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGATTTTAAAAATGGTGTCTG	2903
QY 1801	TTCTACGTGGAAGTCACTGCTCCACAAACCCCACTCAACTCAAGTTTCAAGCTCATC	1860
DB 2904	TTCTACGTGGAAGTCACTGCTCCACAAACCCCACTCAACTCAAGTTTCAAGCTCATC	2963
QY 1861	ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCCCAAGCACAATTAATAGTAGGAC	1920
DB 2964	ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCCCAAGCACAATTAATAGTAGGAC	3023

QY	1921	TACCAATCCTAACAAATAAGCACCAACAATCAAATCAACCCCGTGCATCAACAAAATCAGAA	1980
Dd	3024	TACCAATCCTAACAAATAAGCACCAACAATCAAATCAACCCCGTGCATCAACAAAATCAGAA	3083
Qy	1981	TCCTCAACCAGCACCAACCA	1999
Dd	3084	TCCTCAACCAGCACCAACCA	3102
RESULT 10 AE008414/c			
LOCUS			
DEFINITION			
Streptococcus pneumoniae R6 section 30 of 184 of the complete genome.			
ACCESSION			
AE008414 AE007317			
VERSION			
AE008414.1 GI:15457885			
KEYWORDS			
SOURCE			
Streptococcus pneumoniae R6.			
ORGANISM			
Streptococcus pneumoniae R6			
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.			
REFERENCE			
1 (bases 1 to 10313)			
AUTHORS	Hoskins,J.A., Alborn,W. Jr., Arnold,J., Blaszcak,L., Burgett,S., DeHoff,B.S., Estrem,S., Fritz,L., Fu,D.-J., Fuller,W., Geringer,C., Gilmour,R., Glass,J.S., Khoja,H., Kraft,A., Lagace,R., LeBlanc,D.J., Lee,L.N., Lefkowitz,E.J., Lu,J., Matsushima,P., McAhren,S., McHenry,M., McLeaster,K., Mundy,C., Niclas,T.I., Norris,F.H., O'Gara,M., Peery,R., Robertson,G.T., Rockey,P., Sun,P.-M., Winkler,M.E., Yang,Y., Young-Bellido,M., Zhao,G., Zook,C., Baltz,R.H., Jaskunas,S.R., Rosteck,P.R. Jr., Skatrud,P.L. and Glass,J.I.	J. Bacteriol.	183 (19), 5709-5717 (2001)
TITLE			
JOURNAL	Genome of the Bacterium Streptococcus pneumoniae Strain R6		
MEDLINE	21429245		
PUBMED	11544234		
REFERENCE	2 (bases 1 to 10313)		
AUTHORS	Hoskins,J.A., Alborn,W. Jr., Arnold,J., Blaszcak,L., Burgett,S., DeHoff,B.S., Estrem,S., Fritz,L., Fu,D.-J., Fuller,W., Geringer,C., Gilmour,R., Glass,J.S., Hann,A., Khoja,H., Kraft,A., LaGace,R., LeBlanc,D.J., Lee,L.N., Lefkowitz,E.J., Lu,J., Matsushima,P., McAhren,S., McHenry,M., McLeaster,K., Mundy,C., Niclas,T.I., Norris,F.H., O'Gara,M., Peery,R., Robertson,G.T., Rockey,P., Sun,P.-M., Winkler,M.E., Yang,Y., Young-Bellido,M., Zhao,G., Zook,C., Baltz,R.H., Jaskunas,S.R., Rosteck,P.R. Jr., Skatrud,P.L. and Glass,J.I.	Direct Submission	
Submitted (27-JUL-2001) Infectious Diseases Research, Eli Lilly and Company, Lilly Research Labs, Indianapolis, IN 46285-0438, USA			
Location/Qualifiers			
1..10313			
/organism="Streptococcus pneumoniae R6"			
/strain="R6"			
/db_xref="taxon:171101"			
complement(114..2273)			
/gene="pbpa"			
/note="spr0329"			
complement(114..2273)			
/gene="pbpa"			
/EC_number="2.3.2."			
/EC_number="2.4.1.129"			
/codon_start=1			
/transl_table=11			
/label=spr0329			
/product="Penicillin-binding protein 1A"			
/protein_id="AAK99133.1"			
/db_xref="GI:15457886"			
/translation="MNKPITLRLIKYLISFLSVIAIIVLGGVFFYYVSKAPSLSE SKLVATTSTKYDNKNOLIADGERVRNAOANDIPTDLVKAIYSIEDHFDFHRGID TRILGALTNRLONSLOGSALTCOLIKLTFTSDOTISRKADEAWLAIQLEOK ATKQEILTYINKYMNGCYGMQTAAQNYYGKDNLNLSLPQLALLAGMPAQNQYDP YSHPBAADRNNLRNVISEMKNOGYISAQYEKAVNPITDGLGSUKASNPAYMDNTL KEVINQAQVEETGYLLTTGMDEVYNVDQEAQHKLWPDINTDDYVAYPDDDELQVAITVL			
REV 1			


```
CDS
/ gene= "spr0334"
5514..6908
/ gene= "spr0334"
/codon_start=1
/transl_table=11
/label=spr0334
/product="Conserved hypothetical protein"
/protein_id="AAK99138.1"
/db_xref="GI:15457891"
/translation="MSKKRRNRHKKEAQPFQDFDEAKELTVQCAIRKNEEVEAGVLP
ESILDKIVKQRHREIDAKFATQKKKEEFVETSLDDLIQEMREAVEKSEASSEEV
PSEDLILPLDEEQGLDPLLDENPTTEEEVEEQNLSRLQDEDSKKSKGF
ILTVLALSVIICVAYVYVQVARSKEIETQSQTANQSDVDDENTLYDAFYNSN
KTKLNSOPDKLSLTKLIDKLEGSREHTLAKSKYDSLATQIKAQDVNAOPEKPAIV
DGLVDTNAKASDAKFTDIKTGNTLKDVKDKAI SLGKSOQTSSTSSSSSOFSSTSSS
OASSWTSEPKSPSSNETRSSSEVMGILSSAGVAVORSARVAYNOSAIDDSNSAM
FRADGVLTQILATSRSGYITGDQILERNVINGNYLNKPKDGTLYLFTLNCKTGY
FVNGAGHADDUDY"
6984..8408
/ gene= "gnd"
/ note= "spr0335"
6984..8408
/ gene= "gnd"
/BC_number="1.1.1.44"
/codon_start=1
/transl_table=11
/label=spr0335
/product="6-phosphogluconate dehydratase"
/protein_id="AAK99139.1"
/db_xref="GI:15457892"
/translation="MTKANFGVVGMAVMGRNIALNIESRGYTVAIYNSRKEKTEDVIA
CHPEKNFVPSDVESFVNSIEKPRIMLVQAPGTDATIQALLPHLDKGDILIDGNN
TPYKDTIRNEELANSNGINFICTGVSGGKEGALGPSIMPGQKEAYELIVADVLEIS
AKAPEDGKPCVTVIGPDGAGHYVKMVHNGIEYGMOLIAESYDLMOHLLGLSAEDMAE
IFTEKNKGELSYLIEITADILSRKDDGQDGPVIDYILDAAGNKGTCKWTSQSLDL
GVPLSLITSEVPARISTYKEERVHASKVLPKPAAFNEGDKAELEKIRQALYSKI
ISYAGQFLRVASKEENWNLFPADIASIWRDGCIIIRFLQKITDAYNRDADLANLL
LDEYFDYTKYQOAVRDIVALAVQGVPTFSAITYFDSYRSADLPANLIQIQRD
YFGAITYQRKDEKGFHYSWYDEK"
8420..9109
/ gene= "csr"
/ note= "spr0336"
8420..9109
/ gene= "csr"
/ note= "Putative"
/codon_start=1
/transl_table=11
/label=spr0336
/product="Response regulator"
/protein_id="AAK99140.1"
/db_xref="GI:15457893"
/translation="MGKRILLLEKERNLAHFLSLELQKQYRVDLVEGQKALSMAIQ
TPDYILLNVNGLMDMAODFAEKLRTKPSVIMILDHWDLEQEEVYVQREFAVSIVY
KPVLIENLVARISALFRGRDFIDQHCSLMKVPTVNLRIDVHEHTVYRGEEMIALTR
REYDLATLILGSKKVLTRQQLLESVWKYESATETNIVDVIYIRLRSKLDVKGQKSYIK
TVRGVYTMQE"
9207..10223
/ gene= "cbp"
/ note= "spr0337"
9207..10223
/ gene= "cbp"
/codon_start=1
/transl_table=11
/label=spr0337
/product="Choline-binding protein F"
/protein_id="AAK99141.1"
/ gene= "gnd"
71.9%; Score 1438; DB 1; Length 10313;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```
Query Match
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1 TAAATCTAGCAATAAAAACTCAATGCTGACTTGGGTTCTGGAACGCGCGTCAA 60
|||||
```

```
Db 2115 TAAATCTAGCAATAAAAACTCAATGCTGACTTGGGTTCTGGAACGCGCGTCAA 2056
QY 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCAAGACCA 120
Db 2055 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCAAGACCA 1996
QY 121 TCGCTTCTTCGACCACAGGGGGATTGATACCATCCGTTATCCTGGAGCTTCTTTCGGCAA 180
Db 1995 TCGCTTCTTCGACCACAGGGGGATTGATACCATCCGTTATCCTGGAGCTTCTTTCGGCAA 1936
QY 181 TCTGAAACAATTCCTCCCAAGGTGATCAACTCTACCCCAACAGTTGATTAAGTTGAC 240
Db 1935 TCTGAAACAATTCCTCCCAAGGTGATCAACTCTACCCCAACAGTTGATTAAGTTGAC 1876
QY 241 TTACTTTTCAACTTCGACTTCCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 1875 TTACTTTTCAACTTCGACTTCCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 1816
QY 301 AGCGATTCAAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
Db 1815 AGCGATTCAAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 1756
QY 361 GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAAACTACTATGTTAA 420
Db 1755 GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAAACTACTATGTTAA 1696
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTTCGCTGGCTGGAATGCTCAGGCACC 480
Db 1695 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTTCGCTGGCTGGAATGCTCAGGCACC 1636
QY 481 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGACCGCGGGAACCTTGTCTT 540
Db 1635 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGACCGCGGGAACCTTGTCTT 1576
QY 541 ATCTGAATGAAATAACAGGCTACATCTCTGCTGACAGATATGAGAAACAGTCAATAC 600
Db 1575 ATCTGAATGAAATAACAGGCTACATCTCTGCTGACAGATATGAGAAACAGTCAATAC 1516
QY 601 ACCAATATCTGATGGACTACAAAGTCTCAAACTCAGCAAGTAATTTACCTCTTACATGGA 660
Db 1515 ACCAATATCTGATGGCTACAAAGTCTCAAACTCAGCAAGTAATTTACCTCTTACATGGA 1456
QY 661 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAACAGGCTATAACCTACTCAC 720
Db 1455 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAACAGGCTATAACCTACTCAC 1396
QY 721 AACTGGGATGGATGCTCTACACAAATGTAGACCAGAAAGCTCAAAACATCTGTGGGATAT 780
Db 1395 AACTGGGATGGATGCTCTACACAAATGTAGACCAGAAAGCTCAAAACATCTGTGGGATAT 1336
QY 781 TTACAATACAGCAATAGTTGCCCTATCCAGACGATGAATTCGAAGTCGCTTCTACCAT 840
Db 1335 TTACAATACAGCAATAGTTGCCCTATCCAGACGATGAATTCGAAGTCGCTTCTACCAT 1276
QY 841 TGTGTATGTTTCTAAGCGTAAAGTCATTCGCCAGCTAGGACGACGCATCAGTCAAGTAA 900
Db 1275 TGTGTATGTTTCTAAGCGTAAAGTCATTCGCCAGCTAGGACGACGCATCAGTCAAGTAA 1216
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAAAACACCGGACTGGGGATCAACTATGAA 960
Db 1215 TGTTCCTTCGGAATTAACCAAGCAGTAGAAAACACCGGACTGGGGATCAACTATGAA 1156
QY 961 ACCGATCACAGACTATGCTCCTGCTGGAGTACGGTGTCTACGATTCAACTGCTACTAT 1020
Db 1155 ACCGATCACAGACTATGCTCCTGCTGGAGTACGGTGTCTACGAGTCAACTGCCACTAT 1096
QY 1021 CGTTACGATGAGCCCTATAACTTACCTGGGACAAATACTCCTGTTTATTAACCTGGGATAG 1080
Db 1095 CGTTACGATGAGCCCTATAACTTACCTGGGACAAATACCCTGTTTATTAACCTGGGATAG 1036
QY 1081 GGCGTACTTTGGCAACATCACCTTGAATAACGCCCTGCAACATCGGAACCTGCCAGC 1140
Db 1035 GGCGTACTTTGGCAACATCACCTTGAATAACGCCCTGCAACATCGGAACCTGCCAGC 976
```



```
QY 1141 CQTGGAAGCTCTAAACAAGTCGGACTCAACCGCGCAAGACTTTCTTAATGGTCTAGG 1200
|||||
Db 975 CQTGGAAGCTCTAAACAAGTCGGACTCAACCGCGCAAGACTTTCTTAATGGTCTCGG 916
|||||
QY 1201 AATCGACTACCAAGTATTTCACCTACTCAATGCGCATTTCAAGTAACACACCGCAATCAGA 1260
|||||
Db 915 AATCGACTACCAAGTATTTCACCTACTCAATGCGCATTTCAAGTAACACACCGCAATCAGA 856
|||||
QY 1261 CAAAAAATATGAGCAAGTAGTGAAGAGATGGCTGCTGTCTAGCTGCGCTTTGCAAAATGG 1320
|||||
Db 855 CAAAAAATATGAGCAAGTAGTGAAGAGATGGCTGCTGTCTAGCTGCGCTTTGCAAAATGG 796
|||||
QY 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTAGTGATGGGAGTGAAGA 1380
|||||
Db 795 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTAGTGATGGGAGTGAAGA 736
|||||
QY 1381 AGAGTCTCTAATGTCGGAAGTCGTCGCATGAAGGAACGACAGCCTATATGATGACCGA 1440
|||||
Db 735 AGAGTCTCTAATGTCGGAAGTCGTCGCATGAAGGAACGACAGCCTATATGATGACCGA 676
|||||
QY 1441 CATGATGAAGACAGCTCTGACCTTATGGAACCTGACGAAATGCTATCTTGTGGCTCCC 1500
|||||
Db 675 CATGATGAAGACAGCTCTGAGTTATGGAACCTGACGAAATGCTATCTTGTGGCTCCC 616
|||||
QY 1501 TCAGGCTGTATAAACAGGAACCTCTAACTATACAGACGAGGAATTTGAAACCCACATCAA 1560
|||||
Db 615 TCAGGCTGTATAAACAGGAACCTCTAACTATACAGACGAGGAATTTGAAACCCACATCAA 556
|||||
QY 1561 GACCTCTCAATTTGTAGCACTGTAGCACTATTTGCTGCTATACCGGTAATATTTCAAT 1620
|||||
Db 555 GACCTCTCAATTTGTAGCACTGTAGCACTATTTGCTGCTATACCGGTAATATTTCAAT 496
|||||
QY 1621 GGCTGTATGACAGGCTATTCTAACCGCTGTACACCACTTTGTAGGCAATGGCTTACGCT 1680
|||||
Db 495 GGCTGTATGACAGGCTATTCTAACCGCTGTACACCACTTTGTAGGCAATGGCTTACGCT 436
|||||
QY 1681 CGCTGCAAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAGA 1740
|||||
Db 435 CGCTGCAAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAGA 376
|||||
QY 1741 TTGGATATACAGAGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 1800
|||||
Db 375 TTGGATATACAGAGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 316
|||||
QY 1801 TTCTAGTGAAGCTCACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
|||||
Db 315 TTCTAGTGAAGCTCACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 256
|||||
QY 1861 ATCAGATAGTTCAACTTCAAGTACAGTACCTCAACCACTCCAAAGCACAATAATAGTACGAC 1920
|||||
Db 255 ATCAGATAGTTCAACTTCAAGTACAGTACCTCAACCACTCCAAAGCACAATAATAGTACGAC 196
|||||
QY 1921 TACCAATCTTACAAATATACCAACATCAATCAATCAACCCCTGATCAACAAATCAGAA 1980
|||||
Db 195 TACCAATCTTACAAATATACCAACATCAATCAATCAACCCCTGATCAACAAATCAGAA 136
|||||
QY 1981 TCCTCAACAGCAGCAACCA 1999
|||||
Db 135 TCCTCAACAGCAGCAACCA 117
|||||

RESULT 11
AX111327
LOCUS AX111327 1249 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 2060 from Patent WO0123604.
ACCESSION AX111327
VERSION AX111327.1 GI:13927619
KEYWORDS
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
```

```
REFERENCE 1 (bases 1 to 1249)
AUTHORS Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 2060 05-APR-2001;
FEATURES Infectio Diagnostic (I.D.I.) INC. (CA)
Location/Qualifiers
source 1. .1249
/organism="Streptococcus pneumoniae"
/strain="StrR-57"
BASE COUNT 402 a 302 c 246 g 299 t
ORIGIN
Query Match 59.9%; Score 1198; DB 6; Length 1249;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1248; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 720 CAACCTGGATGGATGCTACACAAATGTACCAAGAGCTCAAAAACATCTGTGGGATA 779
|||||
Db 1 CAACCTGGATGGATGCTACACAAATGTACCAAGAGCTCAAAAACATCTGTGGGATA 60
|||||
QY 780 TTTTACAATACAGACGAATAGTTGCTTATCCAGACGATGAATTGCAAGTCGCTTACCA 839
|||||
Db 61 TTTTACAATACAGACGAATAGTTGCTTATCCAGACGATGAATTGCAAGTCGCTTACCA 120
|||||
QY 840 TTGTTGATTTTCTAAGGTTAAAGTCATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTA 899
|||||
Db 121 TTGTTGATTTTCTAAGGTTAAAGTCATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTA 180
|||||
QY 900 ATGTTTCTTCGGAATTAACCAAGCAGTAGTAACAACCCGCTGGGGATCAACTATGA 959
|||||
Db 181 ATGTTTCTTCGGAATTAACCAAGCAGTAGTAACAACCCGCTGGGGATCAACTATGA 240
|||||
QY 960 AACCGATCAGACTATGCTTCCTGCTTGAGTACGGTGTCTACGATTCGAATTCGCTACTA 1019
|||||
Db 241 AACCGATCAGACTATGCTTCCTGCTTGAGTACGGTGTCTACGATTCGAATTCGCTACTA 300
|||||
QY 1020 TCCTTCCAGATGAGCCCTATAACTACCTGGGACAATACTCCTGTTTATAACTGGGATA 1079
|||||
Db 301 TCCTTCCAGATGAGCCCTATAACTACCTGGGACAATACTCCTGTTTATAACTGGGATA 360
|||||
QY 1080 GGGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAATCGGAAACGTCGCCAG 1139
|||||
Db 361 GGGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAATCGGAAACGTCGCCAG 420
|||||
QY 1140 CCGTGAAACTCTAAACAAGTCGGACTCAACCGCGCAAGACTTTCCCTAAATGGTCTAG 1199
|||||
Db 421 CCGTGAAACTCTAAACAAGTCGGACTCAACCGCGCAAGACTTTCCCTAAATGGTCTAG 480
|||||
QY 1200 GAATCAGTACCAACTATTCACCTACTCAATGCCATTTCAAGTAACACACCAATCAG 1259
|||||
Db 481 GAATCAGTACCAACTATTCACCTACTCAATGCCATTTCAAGTAACACACCAATCAG 540
|||||
QY 1260 AAAAAAATATGAGCAAGTAGTGAAGAGATGCTGCTTACGCTGCTTTGCAAAATG 1319
|||||
Db 541 AAAAAAATATGAGCAAGTAGTGAAGAGATGCTGCTTACGCTGCTTTGCAAAATG 600
|||||
QY 1320 GTGGAACTTACTATAAACCAATGTATATCAATAAAGTCGCTTTAGTATGGAGTGAAG 1379
|||||
Db 601 GTGGAACTTACTATAAACCAATGTATATCAATAAAGTCGCTTTAGTATGGAGTGAAG 660
|||||
QY 1380 AAGAGTTCTTAATGTCGGAACCTCGGCATGAAGGAAGCAGCCTATATGATGACCG 1439
|||||
Db 661 AAGAGTTCTTAATGTCGGAACCTCGGCATGAAGGAAGCAGCCTATATGATGACCG 720
|||||
QY 1440 ACATGATGAAGACAGTCTTGACTTATGGAACCTGCAAGAAATGCTATCTTGTGCTGCC 1499
|||||
Db 721 ACATGATGAAGACAGTCTTGACTTATGGAACCTGCAAGAAATGCTATCTTGTGCTGCC 780
|||||
QY 1500 CTCAGGCTGGTAAACAGGAACCTCTAATATACAGAGGAATTTGAAACCAACATCA 1559
|||||
```


|||||
Db 781 CTCAGCGTGGTAAACAGAACTCTAACTATACAGACGAGGAAATGAAAACCATCA 840
QY 1560 AGACCTCTCAATTTGAGCACCTGATGAACCTATTTGCTGGCTATACGCTTAATATTCAA 1619
Db 841 AGACCTCTCAATTTGAGCACCTGATGAACCTATTTGCTGGCTATACGCTTAATATTCAA 900
QY 1620 TGGCTGATGGACAGGCTATTTCTAACCGTCTGACAGCACTTTGAGCAATGGCCTTACGG 1679
Db 901 TGGCTGATGGACAGGCTATTTCTAACCGTCTGACAGCACTTTGAGCAATGGCCTTACGG 960
QY 1680 TGGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTCTGAAGGAACAATFCCAGAAG 1739
Db 961 TGGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTCTGAAGGAACAATFCCAGAAG 1020
QY 1740 ATTGGAAATATACAGAGGGCTCTACAGAAATGGAGAATTCGTATTTAAAAATGGTGCTC 1799
Db 1021 ATTGGAAATATACAGAGGGCTCTACAGAAATGGAGAATTCGTATTTAAAAATGGTGCTC 1080
QY 1800 GTTCTACGTGGAACTCACCTGCTCCACAAACCCCACTCAACTGAAAGTTCAAGCTCAT 1859
Db 1081 GTTCTACGTGGAGCTCACCTGCTCCACAAACCCCACTCAACTGAAAGTTCAAGCTCAT 1140
QY 1860 CATCAGATAGTTCAACTTACAGCTCTAGCTCAACCACTCCAGCACAATAATAGTACGA 1919
Db 1141 CATCAGATAGTTCAACTTACAGCTCTAGCTCAACCACTCCAGCACAATAATAGTACGA 1200
QY 1920 CTACCAATCCTTACCAATAATACGCAACAATCAAAATCAACCCCTGATCA 1968
Db 1201 CTACCAATCCTTACCAATAATACGCAACAATCAAAATCAACCCCTGATCA 1249

RESULT 12
AB006874
LOCUS AB006874 1195 bp DNA linear BCT 13-FEB-1999
DEFINITION Streptococcus pneumoniae pbpla gene for penicillin binding protein 1A, partial cds, isolate:#10/219.
ACCESSION AB006874
VERSION AB006874.1 GI:3395651
KEYWORDS pbpla; penicillin binding protein 1A.
SOURCE Streptococcus pneumoniae (isolate:#10/219) DNA.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.
REFERENCE 1 (bases 1 to 1195)
AUTHORS Asahi,Y.
TITLE Direct Submission
JOURNAL Submitted (27-AUG-1997) Yasuko Asahi, Teikyo University School of Medicine, Department of Clinical Pathology; 2-11-1 Kaga, Itabashi-ku, Tokyo 173-8605, Japan (E-mail:asahi@med.teikyo-u.ac.jp, Tel:81-3-3964-1211, Fax:81-3-3963-6023)
2 (sites)
REFERENCE Asahi,Y. and Ubukata,K.
AUTHORS Association of a thr-371 substitution in a conserved amino acid motif of penicillin-binding protein 1A with penicillin resistance of Streptococcus pneumoniae
TITLE Antimicrob. Agents Chemother. 42 (9), 2267-2273 (1998)
JOURNAL 98409715
MEDLINE Location/Qualifiers
FEATURES
1. .1195
/organism="Streptococcus pneumoniae"
/isolate="#10/219"
/db_xref="taxon:1313"
1. .1195
/gene="pbpla"
<1. .>1195
/gene="pbpla"
/codon_start=1
/transl_table=11
/product="penicillin binding protein 1A"
/protein_id="BAA32070.1"
/db_xref="GI:3395652"

/translation="VAYPDELOQVASTIVDSNGKVIQALGARHOSNSYFSGINQAVE
TNRDWSMKPITDYAPALEYGVYDSTATIVHDEPNYNTPGTNTPVYNDRGVFGNITL
OYALQOSNRNPAVETLNKVLNKRATFLNGLIDYPSIHYNSAINTTAYMDESKKYGAS
SEKMAAAYAAAFANGCTYKPKYIHKVFSDDGSEKESFNSVGTTRAMKETAYMTDMKT
VLISYGTNRNAYLAWLPQAGKTYSNTDEIEIENHKTQSOFVAPDELAFAGYTRKYSNAV
WTGYSNRLLTPLVGNLTVAAKYRSMYTLSEGSNPEDWNIPEGLYRNGEVEFKNGAR
STWNSPAPQPPSTESSSSSDSSTQSSTSTPSTNNSTTDPNNTTQOQNTTDPQON
ONPOPA"
BASE COUNT 381 a 297 c 233 g 284 t
ORIGIN
Query Match 44.5%; Score 889; DB 1; Length 1195;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1189; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 800 GTTGCCCTATCCAGACATGAATTCGAAGTCGCTTCTACCATGCTTGTATGCTTCTACGGT 859
Db 1 GTTGCCCTATCCAGACATGAATTCGAAGTCGCTTCTACCATGCTTGTATGCTTCTACGGT.60
QY 860 AAAGTCATTGGCCAGCTAGGAGCAGCCCATCAGTCAAGTAATGTTTCTTCGGAATTAAC 919
Db 61 AAAGTCATTGGCCAGCTAGGAGCAGCCCATCAGTCAAGTAATGTTTCTTCGGAATTAAC 120
QY 920 CAAGCAGTAGAACAACCCGCGACTGGGATCAACTATGAAACCGATCAGACTATGCT 979
Db 121 CAAGCAGTAGAACAACCCGCGACTGGGATCAACTATGAAACCGATCAGACTATGCT 180
QY 980 CTTGCCCTTGGAGTACGGTGTCTACGATTCAACTGCTACTATCGTTCAAGTACGCCCTAT 1039
Db 181 CTTGCCCTTGGAGTACGGTGTCTACGATTCAACTGCTACTATCGTTCAAGTACGCCCTAT 240
QY 1040 AACTACCCCTGGGACAATCTCTCTGTTTAACTGGGATAGGGCTACTTTTGCAACATC 1099
Db 241 AACTACCCCTGGGACAATCTCTCTGTTTAACTGGGATAGGGCTACTTTTGCAACATC 300
QY 1100 ACCTTGCAATACGCCCTGCAACAATCGGAAACGTCGCCGCTGGAAACTCTTAAACAAG 1159
Db 301 ACCTTGCAATACGCCCTGCAACAATCGGAAACGTCGCCGCTGGAAACTCTTAAACAAG 360
QY 1160 GTCGGACTCAACCGCGCCGAAGACTTTCTTAATGGTCTAGGAATCGACTACCAAGTATT 1219
Db 361 GTCGGACTCAACCGCGCCGAAGACTTTCTTAATGGTCTAGGAATCGACTACCAAGTATT 420
QY 1220 CACTACTCAAAATGCCATTTCAAGTAACACAAACCGAATCAGACAAAATATGGAGCAAGT 1279
Db 421 CACTACTCAAAATGCCATTTCAAGTAACACACACGGAATCAGACAAAATATGGAGCAAGT 480
QY 1280 AGTGAAGATGGCTGCTGCTTACGCTGCTTTGCAAAATGGTGGAACTTACTATAACCA 1339
Db 481 AGTGAAGATGGCTGCTGCTTACGCTGCTTTGCAAAATGGTGGAACTTACTATAACCA 540
QY 1340 ATGTATATCCATAAAGTCGCTTTTACTGATGGAGTGAAAAGAGTCTCTTAATGTCGA 1399
Db 541 ATGTATATCCATAAAGTCGCTTTTACTGATGGAGTGAAAAGAGTCTCTTAATGTCGA 600
QY 1400 ACTCGTGCCATGAAGGAAACGACAGCTATATGATCAGCAGATGATGAAAACAGTCTTG 1459
Db 601 ACTCGTGCCATGAAGGAAACGACAGCTATATGATCAGCAGATGATGAAAACAGTCTTG 660
QY 1460 ACTTATGGAACGGGCAAAATGCCTATCTTGTGCTGCCCTCCCTCAGGCTGATAAACAGGA 1519
Db 661 AGTTATGGAACGGGCAAAATGCCTATCTTGTGCTGCCCTCCCTCAGGCTGATAAACAGGA 720
QY 1520 ACCTCTAACTATACAGCAGGAAATGAAAACCATCAAGACCTCTCAATTTGTAGCA 1579
Db 721 ACCTCTAACTATACAGCAGGAAATGAAAACCATCAAGACCTCTCAATTTGTAGCA 780
QY 1580 CTTGATGAACCTATTTGCTGCTATACGCTTAATATTTCAATGGCTGTATGGACAGGCTAT 1639
Db 781 CTTGACGAACCTATTTGCTGCTATACGCTTAATATTTCAATGGCTGTATGGACAGGCTAT 840
QY 1640 TCTAACCGCTGTGACACCACTTGTAGGCAATGGCTTACGGTCTGCTGCCAAAGTTTACCGC 1699

|||||
Db 841 TCTAACCGTCTGACACCACTTTAGGCAATGGCTTACGGTGGCTGCCAAAGTTTACCGC 900
QY 1700 TCTATGATGACCTACTCTCTCAAGGACCAATCCAGAGATTGGAATATACAGAGGG 1759
Db 901 TCTATGATGACCTACTCTCTCAAGGACCAATCCAGAGATTGGAATATACAGAGGG 960
QY 1760 CTCTACAGAAATGGAGATTCTGATTTAAATGGTCTGCTTCTACGTGGAACCTCACCT 1819
Db 961 CTCTACAGAAATGGAGATTCTGATTTAAATGGTCTGCTTCTACGTGGAACCTCACCT 1020
QY 1820 GCTCACAAACACCCCACTCAACTGAAGTTCAAGTCTCATCATCATCATCATCATCAT 1879
Db 1021 GCTCACAAACACCCCACTCAACTGAAGTTCAAGTCTCATCATCATCATCATCATCAT 1080
QY 1880 CAGTCTAGCTCAACCACTCCAGGACCAATATATAGTACGACTACCAATCTCAACAATAT 1939
Db 1081 CAGTCTAGCTCAACCACTCCAGGACCAATATATAGTACGACTACCAATCTCAACAATAT 1140
QY 1940 ACGCAACCAATCAATACACCCCTGATCAACAAATCAAGATCTCAACAGCAC 1994
Db 1141 ACGCAACCAATCAATACACCCCTGATCAACAAATCAAGATCTCAACAGCAC 1195

RESULT 13
AB006868
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AB006868 1195 bp DNA linear BCT 13-FEB-1999
Streptococcus pneumoniae pbpla gene for penicillin binding protein
1A, partial cds, isolate:#1/H23.
AB006868
AB006868.1 GI:3395639
pbpla; penicillin binding protein 1A.
Streptococcus pneumoniae (isolate:#1/H23), DNA.
Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 1195)
Asahi,Y.
Direct Submission
Submitted (27-AUG-1997) Yasuko Asahi, Teikyo University School of
Medicine, Department of Clinical Pathology; 2-11-1 Kaga,
Itabashi-ku, Tokyo 173-8605, Japan
(E-mail:asahi@med.teikyo-u.ac.jp, Tel:81-3-3964-1211,
Fax:81-3-3963-6023)
2 (sites)
Asahi,Y. and Ubukata,K.
Association of a thr-371 substitution in a conserved amino acid
motif of penicillin-binding protein 1A with penicillin resistance
of Streptococcus pneumoniae
Antimicrob. Agents Chemother. 42 (9), 2267-2273 (1998)
98409715

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
FEATURES

Location/Qualifiers
1. 1195
/organism="Streptococcus pneumoniae"
/isolate="#1/H23"
/db_xref="taxon:1313"
1. 1195
/gene="pbpla"
<1. >1195
/gene="pbpla"
/codon_start=1
/transl_table=11
/product="penicillin binding protein 1A"
/protein_id="BA332064.1"
/db_xref="GI:3395640"
/translation="VAIPDDELQVASTIVDSNGKVIQLGARHOSNVSGINQAVE
TNRDWMGSTMKPTIDVPALEYGVYDSTATIVHDEPYNYPGTNTVYNWDRGYFGNITL
QYALQQSRNPVATLNRKVLNRKATFLGLIDYPSIHYNSNAISSNTESDCKKYGAS
SEKWAAYAFANGTYYKPMY IHKVVFDGSKOEFNSVGRFRAMKETAYVNMWDMKT
VLSYGTGRNAYLAWLPQAGKTGNTYDEIENHIKTSOFVAPDELEFAGYTRKYSMV
WTGYSNRLTPLVGNGLVAAKYVRSMYTLSEGSNPEDWNIPEGLIRNGEYFVRNGAR
STWSSPAPQPPSTESSSSSDSSTSSSTTPSTNNSTTTNNPNNNTQQSNTTPDQON
ONPQPA"

BASE COUNT 380 a 299 c 234 g 282 t
ORIGIN
Query Match 41.8%; Score 835; DB 1; Length 1195;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 803 GCTATCCAGAGCATCAATTTGCAAGTCGCTTCTACCATTTGTTGATGCTTTCTAAGGTA 862
Db 4 GCTATCCAGAGCATGAATTTGCAAGTCGCTTCTACCATTTGTTGATGCTTTCTAAGGTA 63
QY 863 GTCATTGCCAGCTAGGAGCAGCCCATCATAGTCAAGTAATGTTTTCCTTCGGAATTAACCA 922
Db 64 GTCATTGCCAGCTAGGAGCAGCCCATCATAGTCAAGTAATGTTTTCCTTCGGAATTAACCA 123
QY 923 GCAGTAGAACAACCGCGACTGGGGATCAACTATGAAACCGGATCACAGACTATGCTCCT 982
Db 124 GCAGTAGAACAACCGCGACTGGGGATCAACTATGAAACCGGATCACAGACTATGCTCCT 183
QY 983 GCTTTGGAGTACGGTGTCTACGATTCAACTTGCTTACTATCGTTCACGATGAGCCCTATAAC 1042
Db 184 GCTTTGGAGTACGGTGTCTACGATTCAACTTGCTTACTATCGTTCACGATGAGCCCTATAAC 243
QY 1043 TACCTGGGACAATACTCCTGTTTATACTGGGATAGGGGCTACTTTGGCAACATCAACC 1102
Db 244 TACCTGGGACAATACTCCTGTTTATACTGGGATAGGGGCTACTTTGGCAACATCAACC 303
QY 1103 TTGCAATAGCCCTGCAACAANTCGGAAACGTCGCCAGCGTGGAAACTCTAAACAAGGTC 1162
Db 304 TTGCAATAGCCCTGCAACAANTCGGAAACGTCGCCAGCGTGGAAACTCTAAACAAGGTC 363
QY 1163 GGACTCAACCGCGCCCAAGACTTTCTTAATGGTCTAGGAATCGACTACCCCAAGTATTAC 1222
Db 364 GGACTCAACCGCGCCCAAGACTTTCTTAATGGTCTAGGAATCGACTACCCCAAGTATTAC 423
QY 1223 TACTCAATGCCATTTCAAAGTAAACAACGAATCAGACAAAATAATATGGAGCAAGTAGT 1282
Db 424 TACTCAATGCCATTTCAAAGTAAACAACGAATCAGACAAAATAATATGGAGCAAGTAGT 483
QY 1283 GAAAGATGGCTGCTGCTTACGCTGCTTTGCAATATGGTGGAACTTACTATATAACCAATG 1342
Db 484 GAAAGATGGCTGCTGCTTACGCTGCTTTGCAATATGGTGGAACTTACTATATAACCAATG 543
QY 1343 TATATCCATAAAGTCTGCTTTAGTGTAGTGGAGTGAAGAGAGTTCTCTAATGCGGAAC 1402
Db 544 TATATCCATAAAGTCTGCTTTAGTGTAGTGGAGTGAAGAGAGTTCTCTAATGCGGAAC 603
QY 1403 CGTGCCATGAAGAAACGACAGCCCTATATATGACCGCATGATGAAACAGTCTTTGACT 1462
Db 604 CGTGCCATGAAGAAACGACAGCCCTATATATGACCGCATGATGAAACAGTCTTTGACT 663
QY 1463 TATGNACTGGAGAAATGCCTATCTTGGTCCCTCAGGCTGGTAAACAGAGAAC 1522
Db 664 TATGNACTGGAGAAATGCCTATCTTGGTCCCTCAGGCTGGTAAACAGAGAAC 723
QY 1523 TCTAATATACAGACGAGAAATTTGAAAACCAATCAAGACCTCTCAATTTGTAGCACCT 1582
Db 724 TCTAATATACAGACGAGAAATTTGAAAACCAATCAAGACCTCTCAATTTGTAGCACCT 783
QY 1583 GATGAACATAATTTGCTGGCTATACGGCTAAATATTCATATGGCTGTATGACAGGCTATTC 1642
Db 784 GACGAACATAATTTGCTGGCTATACGGCTAAATATTCATATGGCTGTATGACAGGCTATTC 843
QY 1643 ACCGCTCTGACACCACTTTAGGCAATGCGCTTACGGCTCGCTGCCAAAGTTTACCGCTCT 1702
Db 844 AACCGCTCTGACACCACTTTAGGCAATGCGCTTACGGCTCGCTGCCAAAGTTTACCGCTCT 903
QY 1703 ATGATGACCTACTCTGCTCAAGGAGCAATCCAGAGATTGGAATATACACAGAGGGGCTC 1762
Db 904 ATGATGACCTACTCTGCTCAAGGAGCAATCCAGAGATTGGAATATACACAGAGGGGCTC 963
QY 1763 TACAGAAATGGAGAAATTCGATTTAAAAATGCTGCTCTCTACGTGGAACCTCACCTGCT 1822


```
|||||
Db 964 TACAGAAATGGAGATTCGTATTAAATGGTGCCTGTACGTGGAGCTCACGTGCT 1023
QY 1823 CCACAAACCCCACTCAACTGAAAGTTCAAGCTCATCATCATGATGTTCAACTTCCACAG 1882
Db 1024 CCACAAACCCCACTCAACTGAAAGTTCAAGCTCATCATCATGATGTTCAACTTCCACAG 1083
QY 1883 TCTAGCTCAACCACTCCAAAGCACAATAATAGTACGACTACCAATCCTTAACAATATAGC 1942
Db 1084 TCTAGCTCAACCACTCCAAAGCACAATAATAGTACGACTACCAATCCTTAACAATATAGC 1143
QY 1943 CAACAAATCAAAATACAAACCCCTGATCAACAAATATAGTACGACTACCAATCCTTAACAAT 1994
Db 1144 CAACAAATCAAAATACAAACCCCTGATCAACAAATATAGTACGACTACCAATCCTTAACAAT 1195

RESULT 14
AR084885
LOCUS AR084885 960 bp DNA linear PAT 01-SEP-2000
DEFINITION Sequence 3 from patent US 5981229.
ACCESSION AR084885
VERSION AR084885.1 GI:10011656
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 960)
AUTHORS Masure H, Robert, Pearce B.J. and Tuomanen, E.
TITLE Bacterial exported proteins and acellular vaccines based thereon
JOURNAL Patent: US 5981229-A 3 09-NOV-1999;
FEATURES Location/Qualifiers
1..960
source /organism="unknown"
BASE COUNT 315 a 224 c 182 g 239 t
ORIGIN

Query Match 39.8%; Score 796; DB 6; Length 960;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATAAATAAATCAACTCATCTGCTGACTGGGTCTGAACGGCGGTCAA 60
Db 12 TAAATCTACGACATAAATAAATCAACTCATCTGCTGACTGGGTCTGAACGGCGGTCAA 71
QY 61 TCGCCAAAGCTAATGATATATCCACAGATTGGTTAGGCAATCGTTTCTATCGAAGACCA 120
Db 72 TCGCCAAAGCTAATGATATATCCACAGATTGGTTAGGCAATCGTTTCTATCGAAGACCA 131
QY 121 TCGCTTCTTCGACACAGGGGATTGATACCATCCGTATCCTGGGAGCTTCTTGGCGAA 180
Db 132 TCGCTTCTTCGACACAGGGGATTGATACCATCCGTATCCTGGGAGCTTCTTGGCGAA 191
QY 181 TCTGCAAGCAATTCCTCCAGGTGGATCACTCTCAACCAACAGTTGATTAAGTTGAC 240
Db 192 TCTGCAAGCAATTCCTCCAGGTGGATCACTCTCAACCAACAGTTGATTAAGTTGAC 251
QY 241 TTACTTTTCACTTCGACATTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGTT 300
Db 252 TTACTTTTCACTTCGACATTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGTT 311
QY 301 ACGGATTACAGTTAGAACAAAAAGCAACCAAGCAAGCAAGAAATCTTGACCTACTATATAATAA 360
Db 312 ACGGATTACAGTTAGAACAAAAAGCAACCAAGCAAGCAAGAAATCTTGACCTACTATATAATAA 371
QY 361 GGTCTACATGCTTAATGGGAATCTGGAATGCAGACAGAGCTCAAACTACTATGGTTAA 420
Db 372 GGTCTACATGCTTAATGGGAATCTGGAATGCAGACAGAGCTCAAACTACTATGGTTAA 431
QY 421 AGACCTCAATAATTTAAGTTTACCTACGTTAGCCTTGCTGGCTGGAATCCCTCAGGCACC 480
Db 432 AGACCTCAATAATTTAAGTTTACCTACGTTAGCCTTGCTGGCTGGAATCCCTCAGGCACC 491
```

```
QY 481 AAACCAATATGACCCCTATTTCATCATCCAGAAGCAGCCCAAGACCGCGGAACTTGGTCTT 540
Db 492 AAACCAATATGACCCCTATTTCATCATCCAGAAGCAGCCCAAGACCGCGGAACTTGGTCTT 551
QY 541 ATCTGAATGAAATCAAGGCTAGATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 552 ATCTGAATGAAATCAAGGCTAGATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 611
QY 601 ACCAATTTACTGATGGACTACAAAGTCTCAATCAGCAAGTAAATTACCTCTGTACATGGA 660
Db 612 ACCAATTTACTGATGGCTACAAGTCTCAATCAGCAAGTAAATTACCTCTGTACATGGA 671
QY 661 TAAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAAGAAACAGGCTATAAACCCTACTCAC 720
Db 672 TAAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAAGAAACAGGCTATAAACCCTACTCAC 731
QY 721 AACTGGGATGGATGCTTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATAT 780
Db 732 AACTGGGATGGATGCTTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATAT 791
QY 781 TTACATACAGAGGAATAGTTCCTTATCCAGACGATGAATTCGAAGTGCCTTCTTACCAT 840
Db 792 TTACATACAGAGGAATAGTTCCTTATCCAGACGATGAATTCGAAGTGCCTTCTTACCAT 851
QY 841 TGTTCATGTTTCTACGGTAAAGTCAATGCCAGCTAGGACGACGCGCATCAGTCAAGTAA 900
Db 852 TGTTCATGTTTCTACGGTAAAGTCAATGCCAGCTAGGACGACGCGCATCAGTCAAGTAA 911
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTGAACAAACCGGACTGGGGA 949
Db 912 TGTTCCTTCGGAATTAACCAAGCAGTGAACAAACCGGACTGGGGA 960

RESULT 15
SPON6
LOCUS S.pneumoniae (8250) ponA gene for penicillin-binding protein 1a.
DEFINITION x67871
ACCESSION X67871.1 GI:47415
VERSION
KEYWORDS penicillin-binding protein 1a; ponA gene.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae.
REFERENCE 1 (bases 1 to 1826)
AUTHORS Hakenbeck R.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek. Genetik, Ihnestr. 73, 1000 Berlin 33, FRG
REFERENCE 2 (bases 1 to 1826)
AUTHORS Martin, C., Sibold, C. and Hakenbeck, R.
TITLE Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain
JOURNAL EMBO J. 11 (11), 3831-3836 (1992)
MEDLINE
FEATURES Location/Qualifiers
1..1826
/organism="Streptococcus pneumoniae"
/strain="8250"
/db_xref="taxon:1313"
/clone="pcMK1-pcMK16"
1..1826
/gene="ponA"
1..>1826
/gene="ponA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="CAA48071.1"
/db_xref="GI:47415"
/db_xref="SPTREMBL:Q54951"
/translation="MNKOTILRIAKYVSICTFLTITAVMLGGGIFLYVVKAPALSE"
```


SKLVATTSKIVYDNKLIADLGSERRVNAQANDIPTDLVKAIVSIEDHRFDHIGID
 TIRILGATLRNLSQSGSTLQOLIKLTYFTSTSDQTSRKAQEAWLAIQLEQK
 ATKQILTYIYINKVYMSNGYGMOTAAQNYGKDLNLSLPOLALLAGPQAPNOYDP
 YSHPEAQRRLNLVSEMKNOYISAEOYKAVNPITDGLQSLKASNPAYMDNYL
 KEVINQVEETCYNLLTGMDOYVYNDQEAOKLWDIYNTDEVVAYPDELQVASTIV
 DYSNGKVAQLGARHQSSNVSGINQAVETNRDWSMKPITDYAPALEYGLGYDSTAT
 IVHDEPNYPGDTPIYINWDRGFYFNITLQYALQOSNRNPVAVETLNKVLGNRAKFTPLN
 GLGIDYPSLSHYNSNAISSNTESDKYIGASSERKMAAAYAAFANGYKPKYIHKVYVES
 DQSEKFEFNVGTFRAMKEATAYMTEMKMTVLTYIGIRGAYLPLWPAQAGTGTSTNYTDE
 EIEKYIKNTGYVADPEMFVGYTRKYVAMAVMTGYSNRLTPIIG*
 BASE COUNT 582 a 422 c 364 g 458 t
 ORIGIN

Query Match 38.7%; Score 774; DB 1; Length 1826;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 924; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATATAAATACTAATTCGTGACTGGGTCTGAACGCCGCGCTCAA 60
 DB 159 TAAATCTACGACATATAAATACTAATTCGTGACTGGGTCTGAACGCCGCGCTCAA 218
 QY 61 TGCCCAAGCTAATGATATATCCACACATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
 DB 219 TGCCCAAGCTAATGATATATCCACACATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
 QY 121 TCGCTTTCTCGACACAGGGGATGATACCATCCGTATCCCTGGGAGCTTCTTGGCGCAA 180
 DB 279 TCGCTTTCTCGACACAGGGGATGATACCATCCGTATCCCTGGGAGCTTCTTGGCGCAA 338
 QY 181 TCTGCAAGCAATTCCTTCAAGGTGGATCACTCTCACCCACAGTTGATTAAGTTGAC 240
 DB 339 TCTGCAAGCAATTCCTTCAAGGTGGATCACTCTCACCCACAGTTGATTAAGTTGAC 398
 QY 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
 DB 399 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458
 QY 301 AGGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGGACCTACTATATAATAA 360
 DB 459 AGGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGGACCTACTATATAATAA 518
 QY 361 GGTCTACATGCTTAATGGGAACATATGGAATGACAGACAGCTCAAACTACTATGTTAA 420
 DB 519 GGTCTACATGCTTAATGGGAACATATGGAATGACAGACAGCTCAAACTACTATGTTAA 578
 QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGTAATGCTCAGGACACC 480
 DB 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGTAATGCTCAGGACACC 638
 QY 481 AAACCAATATGACCCCTATTACATATCCAGAAGCAGCCCAAGACGCCGGAACCTTGGCTTT 540
 DB 639 AAACCAATATGACCCCTATTACATATCCAGAAGCAGCCCAAGACGCCGGAACCTTGGCTTT 698
 QY 541 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGCTCAATAC 600
 DB 699 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGCTCAATAC 758
 QY 601 ACCAATTACTGTGGGACTACAAAGTCTCAAACTCAGCAAGTAATTTACCTGCTTACATGGA 660
 DB 759 ACCAATTACTGTGGGACTACAAAGTCTCAAACTCAGCAAGTAATTTACCTGCTTACATGGA 818
 QY 661 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGAAACAGGCTATAAACCCTACTCAC 720
 DB 819 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGAAACAGGCTATAAACCCTACTCAC 878
 QY 721 AACTGGGATGGATGCTTACAAATGCTAGACCAAGCTCAAAACATCTGTGGGATAT 780
 DB 879 AACTGGGATGGATGCTTACAAATGCTAGACCAAGCTCAAAACATCTGTGGGATAT 938
 QY 781 TTACATATACAGCAATAGTTGCCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 840
 DB 939 TTACATATACAGCAATAGTTGCCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 998

QY 841 TGTGATGTTTCTAACGGTAAAGTCATTGCCAGCTAGGACGCGCATCAGTCAAGTAA 900
 DB 999 TGTGATGTTTCTAACGGTAAAGTCATTGCCAGCTAGGACGCGCATCAGTCAAGTAA 1058
 QY 901 TGTTCCTCTCGGAATTAACCAACAGT 927
 DB 1059 TGTTCCTCTCGGAATTAACCAACAGT 1085

RESULT 16
 AF210748 2154 bp DNA linear BCT 17-SEP-2001
 LOCUS Streptococcus pneumoniae isolate SP 1258 penicillin-binding protein
 DEFINITION 1A (pbpla) gene, partial cds.
 ACCESSION AF210748
 VERSION AF210748.1 GI:563342
 KEYWORDS Streptococcus pneumoniae.
 SOURCE Streptococcus pneumoniae.
 ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group: Streptococcaceae;
 Streptococcus.
 REFERENCE 1 (bases 1 to 2154)
 AUTHORS Ferroni,A. and Berche,P.
 TITLE Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
 penicillin-resistant clinical isolates of Streptococcus pneumoniae
 serotype 23F from the nasopharyngeal flora of children
 J. Med. Microbiol. 50 (9), 828-832 (2001)
 MEDLINE 21432820
 PUBMED 11549185
 REFERENCE 2 (bases 1 to 2154)
 AUTHORS Ferroni,A. and Berche,P.
 TITLE Direct Submission
 JOURNAL Submitted (03-DEC-1999) Microbiology, Hospital
 Necker-Enfants-Malades, 149 rue de Sevres, Paris 75015, France
 FEATURES
 source
 1..2154
 /organism="Streptococcus pneumoniae"
 /serotype="23F"
 /isolate="SP 1258"
 /db_xref="taxon:1313"
 1..>2154
 /gene="pbpla"
 1..>2154
 /gene="pbpla"
 /codon_start=1
 /transl_table=11
 /product="penicillin-binding protein 1A"
 /protein_id="AAF17258.1"
 /db_xref="GI:563342"

gene
 CDS
 1..>2154
 /gene="pbpla"
 /codon_start=1
 /transl_table=11
 /product="penicillin-binding protein 1A"
 /protein_id="AAF17258.1"
 /db_xref="GI:563342"

BASE COUNT 678 a 497 c 419 g 559 t
 ORIGIN 1 others

Query Match 35.3%; Score 705; DB 1; Length 2154;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 855; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATATAAATACTAATTCGTGACTGGGTCTGAACGCCGCGCTCAA 60
 DB 159 TAAATCTACGACATATAAATACTAATTCGTGACTGGGTCTGAACGCCGCGCTCAA 218

Db	159	TAAATCTACAGCAATAAATAATCAACTATTGCTGACTTGGGTTCTGAACCCCGGTCAA	218	TITLE	Alterations in PBP 1A resistance in Streptococcus pneumoniae Antimicrob. Agents Chemother. 42 (6), 1329-1333 (1998)
Qy	61	TGCCCAAGCTAATGATATTCACAGAGATTTGGTTAAGCAATCGTTTCTATCGAAGACCA	120	JOURNAL	98287565
Db	219	TGCCCAAGCTAATGATATTCACAGAGATTTGGTTAAGCAATCGTTTCTATCGAAGACCA	278	MEDLINE	2 (bases 1 to 930)
Qy	121	TCGCTTCTTCGACCACAGAGGGGATGTATACCATCCGATCCTCGGAGCTTCTTTCGCGAA	180	REFERENCE	Smith, A.M.
Db	279	TCGCTTCTTCGACCACAGAGGGGATGTATACCATCCGATCCTCGGAGCTTCTTTCGCGAA	338	AUTHORS	Direct Submission
Qy	181	TCTCAAGCAATTCCTCCAGCTGGATCAACTCTCACCCAAACAGTTGATTAAGTTGAC	240	TITLE	Submitted (06-FEB-1998) Microbiology, SAIMR Hospital, Johannesburg, Gauteng 2000, South Africa
Db	339	TCTCAAGCAATTCCTCCAGCTGGATCAACTCTCACCCAAACAGTTGATTAAGTTGAC	398	JOURNAL	Location/Qualifiers
Qy	241	TTACTTTTCAACTTCGACTTCGACACAGATTAATTCGTAAGGCTCAGGAAGCTTGGTT	300	FEATURES	source
Db	399	TTACTTTTCAACTTCGACTTCGACACAGATTAATTCGTAAGGCTCAGGAAGCTTGGTT	458	gene	1..>930
Qy	301	AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	360	CDS	/organism="Streptococcus pneumoniae"
Db	459	AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	518		/strain="64429"
Qy	361	GSTCTACATGCTAATGGGAATCTGAATCGACACAGCAGCTCAAAACTACTATGCTAA	420		/db_xref="taxon:1313"
Db	519	GSTCTACATGCTAATGGGAATCTGAATCGACACAGCAGCTCAAAACTACTATGCTAA	578		<1..>930
Qy	421	AGACCTCAATAATTAAGTTTACCTTCAGTTAGCCTTGTGCTGCTGCAATTCGCTCAGGCACC	480		/gene="ponA"
Db	579	AGACCTCAATAATTAAGTTTACCTTCAGTTAGCCTTGTGCTGCTGCAATTCGCTCAGGCACC	638		<1..>930
Qy	481	AAACCAATATGACCCCTATTTCACATCCAGAGAGCCCAAGACCGCGGAAACTTTGGTCTT	540		/gene="ponA"
Db	639	AAACCAATATGACCCCTATTTCACATCCAGAGAGCCCAAGACCGCGGAAACTTTGGTCTT	698		/note="PBP 1A"
Qy	541	ATCTGAATGAATAATCAAGCTCATCTCTGCTGACAGTATGAGAAAGCAGTCAATAC	600		/codon_start=1
Db	699	ATCTGAATGAATAATCAAGCTCATCTCTGCTGACAGTATGAGAAAGCAGTCAATAC	758		/transl_table=1
Qy	601	ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAAGTAAATPACCCGCTTACATGGA	660		/product="penicillin-binding protein 1A"
Db	759	ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAAGTAAATPACCCGCTTACATGGA	818		/protein_id="AAC24698.1"
Qy	661	TAATTACTCTCAAGGAAGTCAATCAAGTTGAAGAGAAACAGGCTATACCTTACTCAC	720		/db_xref="GI:2911412"
Db	819	TAATTACTCTCAAGGAAGTCAATCAAGTTGAAGAGAAACAGGCTATACCTTACTCAC	878		/translation="LMDIYNTDEYVAYPDEQLQVATIVDVSNKGKVIQALGARHQSSNVSFGINQAVETNRDWSMTKPTIDYAPALEYGVYDSTATIVHDEPNYPTNTPVVNTDRGYFCNITLQYALQOSRNVPAVETLNKVLNRAKTFNLGLGIDYPSIHYSNAINSTTDSKKYGAASSEKMAAAYAFANGGTYYPKMVIHKVVFSDGSEKEFSNVGTRAMKETTAYMTEMKTVLVYIGIRGAYLPMLPQAGKTGTSNYTDEIEKYIKNTGYVAPDEMFEV
Qy	721	AACGTGGATGGATCTTACACAAATGTAGACCAGAGAGCTCAAAAACATCTGTGGGATAT	780	BASE COUNT	288 a 215 c 198 g 229 t
Db	879	AACGTGGATGGATCTTACACAAATGTAGACCAGAGAGCTCAAAAACATCTGTGGGATAT	938	ORIGIN	
Qy	781	TTACAATACAGCAATACGTTGCTTATCCAGACGATGAATTCGAAGTCGCTTCTACCAT	840	Query Match	33.6%; Score 671; DB 1; Length 930;
Db	939	TTACAATACAGCAATACGTTGCTTATCCAGACGATGAATTCGAAGTCGCTTCTACCAT	998	Best Local Similarity	100.0%; Pred. No. 0;
Qy	841	TGTTGATGTTTCTAACCG 858		Matches	671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	999	TGTTGATGTTTCTAACCG 1016		Qy 770	CTGTGGGATATTTACAATACAGACGAATACGTTGCGCTTATCCAGACGATGAATTCGCAAGTC 829
RESULT 17				Db 1	CTGTGGGATATTTACAATACAGACGAATACGTTGCGCTTATCCAGACGATGAATTCGCAAGTC 60
AF046233				Qy 830	GCTTCTACCAATGTTGATGTTTCTTACCGTAAAGTCAATTCGCCAGCTTAGGACGCGCAT 889
LOCUS				Db 61	GCTTCTACCAATGTTGATGTTTCTTACCGTAAAGTCAATTCGCCAGCTTAGGACGCGCAT 120
DEFINITION				Qy 890	CAGTCAAGTAAATGTTTCCITTCGGAAATTAACAACAGCAGTAAACAACACCGGACTGGGA 949
(ponA) gene, partial cds.				Db 121	CAGTCAAGTAAATGTTTCCITTCGGAAATTAACAACAGCAGTAAACAACACCGGACTGGGA 180
AF046233				Qy 950	TCAACTATGAACCGATCAGACACTATGCTCCTGCTTGGAGTAGGTTGCTACGATTCA 1009
AF046233.1				Db 181	TCAACTATGAACCGATCAGACACTATGCTCCTGCTTGGAGTAGGTTGCTACGATTCA 240
GI:2911411				Qy 1010	ACTGCTACTATCGTTTACGATGAGCCCTTATAACTTACCTTGGGACAAATACTCTCTGTTAT 1069
Streptococcus pneumoniae.				Db 241	ACTGCTACTATCGTTTACGATGAGCCCTTATAACTTACCTTGGGACAAATACTCTCTGTTAT 300
Streptococcus pneumoniae				Qy 1070	AACGTGGATAGGGCTTACTTTGGCAACATCACCTTGGCAATACGCCCTTGAACAATCCGGA 1129
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;				Db 301	AACGTGGATAGGGCTTACTTTGGCAACATCACCTTGGCAATACGCCCTTGAACAATCCGGA 360
Streptococcus.				Qy 1130	AACGTCCACCGCTGGAACTCTAAACAAGTCGGACTCAACCGCGCAAGACTTTCCTA 1189
1 (bases 1 to 930)				Db 361	AACGTCCACCGCTGGAACTCTAAACAAGTCGGACTCAACCGCGCAAGACTTTCCTA 420
Smith, A.M. and Klugman, K.P.				Qy 1190	AATGGTCTAGGAATTCGACTACCCAAAGTATTCTACTTCAATGCCATTTCAAGTAAACACA 1249
				Db 421	AATGGTCTAGGAATTCGACTACCCAAAGTATTCTACTTCAATGCCATTTCAAGTAAACACA 480
				Qy 1250	ACCGAATCAGACAAAAAATATGGAGCAAGTAGTGAAAGATGGCTGCTGTACGCTGCC 1309
				Db 481	ACCGAATCAGACAAAAAATATGGAGCAAGTAGTGAAAGATGGCTGCTGTACGCTGCC 540


```
QY 1310 TTTGCAAAATGGTGAACCTTCTTAATCTCGGAACCTCGTGCATGAAGAAACGACACGCTAT 1429
|||||
Db 541 TTTGCAAAATGGTGAACCTTCTTAATCTCGGAACCTCGTGCATGAAGAAACGACACGCTAT 600
|||||
QY 1370 GGGAGTGAAGAAAGAGTTCTTAATCTCGGAACCTCGTGCATGAAGAAACGACACGCTAT 1429
|||||
Db 601 GGGAGTGAAGAAAGAGTTCTTAATCTCGGAACCTCGTGCATGAAGAAACGACACGCTAT 660
|||||
QY 1430 ATGATGACCGA 1440
|||||
Db 661 ATGATGACCGA 671
|||||

RESULT 18
AF046235 930 bp DNA linear BCT 30-JUN-1998
LOCUS Streptococcus pneumoniae strain 7851 penicillin-binding protein 1A
DEFINITION (ponA) gene, partial cds.
ACCESSION AF046235
VERSION AF046235.1 GI:2911415
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae.
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 930)
AUTHORS Smith,A.M. and Klugman,K.P.
TITLE Alterations in PBP 1A essential for high-level penicillin
resistance in Streptococcus pneumoniae
JOURNAL Antimicrob. Agents Chemother. 42 (6), 1329-1333 (1998)
MEDLINE 98287565
REFERENCE 2 (bases 1 to 930)
AUTHORS Smith,A.M.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1998) Microbiology, SAIMR Hospital, Johannesburg,
Gauteng 2000, South Africa
FEATURES
source
1. .930
Location/Qualifiers
/organism="Streptococcus pneumoniae"
/strain="7851"
/db_xref="taxon:1313"
<1. .>930
/gene="ponA"
<1. .>930
/gene="ponA"
/feature="PBP 1A"
/codon_start=1
/transl_table=1
/product="penicillin-binding protein 1A"
/protein_id="AAC24700.1"
/db_xref="GI:2911416"
/transl_table=1
/translation="LWDIYNTDEVVAYPDDELQVASTIVDSNGKVIQAQLGARHQSSN
VSGINQAVETNRDWSGTMKPTIDYFALEYGVYDSTATIVHDEPNYPTGNTPVYNN
DRGYGNTLOYALQQRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHSNLAISNT
TESDKYTGASERMAAAYAFANGTYIKPMYIKHVFSDGSEKFSNVGVRAMKETT
AYMTEMKTVLSYGTGRNAVLAWLPAQGTGTGNTYDEETENIKNTGYVAPDETFF
GYTRKYSMAVWTGYSNRLTPVGGGFVAAKVYR"
BASE COUNT 283 a 214 c 198 g 235 t
ORIGIN

Query Match 31.0%; Score 620; DB 1; Length 930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 670; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 770 CTGTGGGATATTACAATACAGCAATACGTTGCCTATCCAGACGATGAATTCGAAGTC 829
|||||
Db 1 CTGTGGGATATTACAATACAGCAATACGTTGCCTATCCAGACGATGAATTCGAAGTC 60
|||||
QY 830 GCTTCTACCATGTTGATGTTCTTAACGGTAAAGTCATTCGCCAGCTAGGACGACGCAT 889
|||||
Db 61 GCTTCTACCATGTTGATGTTCTTAACGGTAAAGTCATTCGCCAGCTAGGACGACGCAT 120
|||||
```

```
QY 890 CAGTCAAGTAATGTTTCTTCGGAATTAACCAAGCAGTAGAACAACCAACCGGACTGGGA 949
|||||
Db 121 CAGTCAAGTAATGTTTCTTCGGAATTAACCAAGCAGTAGAACAACCAACCGGACTGGGA 180
|||||
QY 950 TCAACTATGAACCCGATCACAGACTATGCTCTGCTTGGAGTACGGTCTCTACGATTCA 1009
|||||
Db 181 TCAACTATGAACCCGATCACAGACTATGCTCTGCTTGGAGTACGGTCTCTACGATTCA 240
|||||
QY 1010 ACTGCTACTATGTTTCCAGCATGAGCCCTATAACTACCTCGGACAAATACTCTGCTTTAT 1069
|||||
Db 241 ACTGCTACTATGTTTCCAGCATGAGCCCTATAACTACCTCGGACAAATACTCTGCTTTAT 300
|||||
QY 1070 AACTGGGATAGGGGCTACTTTTGGCAACATCACCTTTCGAATACGCCCTGCAACAATGCGA 1129
|||||
Db 301 AACTGGGATAGGGGCTACTTTTGGCAACATCACCTTTCGAATACGCCCTGCAACAATGCGA 360
|||||
QY 1130 AACTGCCAGCCGCTGAAACTCTAAACAGTCCGACTCAACCGCCCAAGACTTTTCCTA 1189
|||||
Db 361 AACTGCCAGCCGCTGAAACTCTAAACAGTCCGACTCAACCGCCCAAGACTTTTCCTA 420
|||||
QY 1190 AATGGTCTAGGAATCGACTACCAAGTATTCACTACTCAATGCCATTTCAAGTAACACA 1249
|||||
Db 421 AATGGTCTAGGAATCGACTACCAAGTATTCACTACTCAATGCCATTTCAAGTAACACA 480
|||||
QY 1250 ACCGAATCAGACAAAAATATATGGAGCAAGTAGTGAAGATGGCTGCTCTTACGCTGCC 1309
|||||
Db 481 ACCGAATCAGACAAAAATATATGGAGCAAGTAGTGAAGATGGCTGCTCTTACGCTGCC 540
|||||
QY 1310 TTTGCAAAATGGTGAACCTTCTTAATCTCGGAACCTCGTGCATGAAGAAACGACACGCTAT 1369
|||||
Db 541 TTTGCAAAATGGTGAACCTTCTTAATCTCGGAACCTCGTGCATGAAGAAACGACACGCTAT 600
|||||
QY 1370 GGGAGTGAAGAAAGAGTTCTTAATCTCGGAACCTCGTGCATGAAGAAACGACACGCTAT 1429
|||||
Db 601 GGGAGTGAAGAAAGAGTTCTTAATCTCGGAACCTCGTGCATGAAGAAACGACACGCTAT 660
|||||
QY 1430 ATGATGACCGA 1440
|||||
Db 661 ATGATGACCGA 671
|||||

RESULT 19
AF046237 930 bp DNA linear BCT 30-JUN-1998
LOCUS Streptococcus pneumoniae strain 17619 penicillin-binding protein 1A
DEFINITION (ponA) gene, partial cds.
ACCESSION AF046237
VERSION AF046237.1 GI:2911419
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae.
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 930)
AUTHORS Smith,A.M. and Klugman,K.P.
TITLE Alterations in PBP 1A essential for high-level penicillin
resistance in Streptococcus pneumoniae
JOURNAL Antimicrob. Agents Chemother. 42 (6), 1329-1333 (1998)
MEDLINE 98287565
REFERENCE 2 (bases 1 to 930)
AUTHORS Smith,A.M.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1998) Microbiology, SAIMR Hospital, Johannesburg,
Gauteng 2000, South Africa
FEATURES
source
1. .930
Location/Qualifiers
/organism="Streptococcus pneumoniae"
/strain="17619"
/db_xref="taxon:1313"
<1. .>930
/gene="ponA"
<1. .>930
/gene="ponA"
/feature="PBP 1A"
/codon_start=1
/transl_table=1
/product="penicillin-binding protein 1A"
/protein_id="AAC24700.1"
/db_xref="GI:2911416"
/transl_table=1
/translation="LWDIYNTDEVVAYPDDELQVASTIVDSNGKVIQAQLGARHQSSN
VSGINQAVETNRDWSGTMKPTIDYFALEYGVYDSTATIVHDEPNYPTGNTPVYNN
DRGYGNTLOYALQQRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHSNLAISNT
TESDKYTGASERMAAAYAFANGTYIKPMYIKHVFSDGSEKFSNVGVRAMKETT
AYMTEMKTVLSYGTGRNAVLAWLPAQGTGTGNTYDEETENIKNTGYVAPDETFF
GYTRKYSMAVWTGYSNRLTPVGGGFVAAKVYR"
BASE COUNT 283 a 214 c 198 g 235 t
ORIGIN
```



```

/gene="pona"
/notes="PBP 1A"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1A"
/protein_id="AAC24702.1"
/db_xref="GI:2911420"
/translation="LWDIYNDEYVAVDPDELOVASTIVDSNGKVIQAOLGARHOSNN
VSFGINQAVETNRDWSMTKIDYAPALEYGVYDSTATIVHDEPYNPYNTVYNN
VDFGFGNTTLOYALQOSRNPVATLKNVGLNRKATFLNGLIDYPSIHYNAISSNT
TESDKKYAKSEKWAAYAAAFANGTYVKPMYIHKVVESDSEKSEFNVGTRAMKETI
AYMTEHKKTVLSVTCGRNAYLAWLPQAKGCTGTSNYDEDEKHKIKNGYVAPDETIV
GYTKRYSNAVMTGYTNRLTPIVIGDFTVAARVYR"
BASE COUNT      289 a 212 c 196 g 233 t
ORIGIN

```

```

Query Match      31.0%; Score 620; DB 1; Length 930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 670; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 770 CTGTGGGATATTACAAATACAGACGAATACGTTGCTATCCAGACGATGAATTGCAAGTC 829
Db 1 CTGTGGGATATTACAAATACAGACGAATACGTTGCTATCCAGACGATGAATTGCAAGTC 60

Qy 830 GCTTCTACCATTTGTTGATGTTTCTAACCGSTAAAGTCATTGCCAGCTAGGAGCAGCCCAT 889
Db 61 GCTTCTACCATTTGTTGATGTTTCTAACCGSTAAAGTCATTGCCAGCTAGGAGCAGCCCAT 120

Qy 890 CAGTCAAGTAATGTTTCTTCGGGAATTAACCAAGCAGTAGAACAACCGCGACGCGGA 949
Db 121 CAGTCAAGTAATGTTTCTTCGGGAATTAACCAAGCAGTAGAACAACCGCGACGCGGA 180

Qy 950 TCAACTATGAAACCGCATCACAGCTATGCTCTGCTGGAGTACGGTCTACGATTCA 1009
Db 181 TCAACTATGAAACCGCATCACAGCTATGCTCTGCTGGAGTACGGTCTACGATTCA 240

Qy 1010 ACTGCTACTATGCTTCACGATGAGCCCTATAACTACCTGGGACAAATACTCCTGTTAT 1069
Db 241 ACTGCTACTATGCTTCACGATGAGCCCTATAACTACCTGGGACAAATACTCCTGTTAT 300

Qy 1070 AACTGGGATAGGGGCTACTTTGGCAACATCACTTTGCAATACGCCCTGCAACAATCGGA 1129
Db 301 AACTGGGATAGGGGCTACTTTGGCAACATCACTTTGCAATACGCCCTGCAACAATCGGA 360

Qy 1130 AACGTCGCCAGCGGTGAAACTCTAACAGAGTCGGACTCAACCGCGCCAAAGACTTTCCCTA 1189
Db 361 AACGTCGCCAGCGGTGAAACTCTAACAGAGTCGGACTCAACCGCGCCAAAGACTTTCCCTA 420

Qy 1190 AATGGTCTAGGAATCGACTACCAAGTATTCATCTACTCAAAATGCCATTTCAAGTAACACA 1249
Db 421 AATGGTCTAGGAATCGACTACCAAGTATTCATCTACTCAAAATGCCATTTCAAGTAACACA 480

Qy 1250 ACCGAATCAGACAAAATAATATGGAGCAAGTAGTGAAGAGATGGCTGCTACCGTGC 1309
Db 481 ACCGAATCAGACAAAATAATATGGAGCAAGTAGTGAAGAGATGGCTGCTACCGTGC 540

Qy 1310 TTTGCAAAATGGTGGAACTTACTATATAACCAATGTATCCATAAAGTCGCTTTAGTGAT 1369
Db 541 TTTGCAAAATGGTGGAACTTACTATATAAACAATGTATCCATAAAGTCGCTTTAGTGAT 600

Qy 1370 GGGAGTGAAGAAAGAGTTCTTAATGTCGGAACCTCGTGCCATGAAGGAAACGACAGCCTAT 1429
Db 601 GGGAGTGAAGAAAGAGTTCTTAATGTCGGAACCTCGTGCCATGAAGGAAACGACAGCCTAT 660

Qy 1430 ATGATGACCGA 1440
Db 661 ATGATGACCGA 671

```

```

RESULT 20
AX111320
LOCUS      930 bp  DNA  linear  PAT 30-APR-2001

```

```

DEFINITION      Sequence 2053 from Patent WO0123604.
ACCESSION      AX111320
VERSION        AX111320.1  GI:13927612
KEYWORDS
SOURCE        Streptococcus pneumoniae.
ORGANISM      Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE      1 (bases 1 to 930)
AUTHORS      Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
Picard,F.J. and Roy,P.H.
TITLE      Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL      Patent: WO 0123604-A 2053 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
FEATURES
Location/Qualifiers
source      1..930
/organism="Streptococcus pneumoniae"
/strain="17619"
/db_xref="taxon:1313"
BASE COUNT      289 a 212 c 196 g 233 t
ORIGIN

```

```

Query Match      31.0%; Score 620; DB 6; Length 930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 670; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 770 CTGTGGGATATTACAAATACAGACGAATACGTTGCTATCCAGACGATGAATTGCAAGTC 829
Db 1 CTGTGGGATATTACAAATACAGACGAATACGTTGCTATCCAGACGATGAATTGCAAGTC 60

Qy 830 GCTTCTACCATTTGTTGATGTTTCTAACCGSTAAAGTCATTGCCAGCTAGGAGCAGCCCAT 889
Db 61 GCTTCTACCATTTGTTGATGTTTCTAACCGSTAAAGTCATTGCCAGCTAGGAGCAGCCCAT 120

Qy 890 CAGTCAAGTAATGTTTCTTCGGGAATTAACCAAGCAGTAGAACAACCGCGACGCGGA 949
Db 121 CAGTCAAGTAATGTTTCTTCGGGAATTAACCAAGCAGTAGAACAACCGCGACGCGGA 180

Qy 950 TCAACTATGAAACCGCATCACAGCTATGCTCTGCTGGAGTACGGTCTACGATTCA 1009
Db 181 TCAACTATGAAACCGCATCACAGCTATGCTCTGCTGGAGTACGGTCTACGATTCA 240

Qy 1010 ACTGCTACTATGCTTCACGATGAGCCCTATAACTACCTGGGACAAATACTCCTGTTAT 1069
Db 241 ACTGCTACTATGCTTCACGATGAGCCCTATAACTACCTGGGACAAATACTCCTGTTAT 300

Qy 1070 AACTGGGATAGGGGCTACTTTGGCAACATCACTTTGCAATACGCCCTGCAACAATCGGA 1129
Db 301 AACTGGGATAGGGGCTACTTTGGCAACATCACTTTGCAATACGCCCTGCAACAATCGGA 360

Qy 1130 AACGTCGCCAGCGGTGAAACTCTAACAGAGTCGGACTCAACCGCGCCAAAGACTTTCCCTA 1189
Db 361 AACGTCGCCAGCGGTGAAACTCTAACAGAGTCGGACTCAACCGCGCCAAAGACTTTCCCTA 420

Qy 1190 AATGGTCTAGGAATCGACTACCAAGTATTCATCTACTCAAAATGCCATTTCAAGTAACACA 1249
Db 421 AATGGTCTAGGAATCGACTACCAAGTATTCATCTACTCAAAATGCCATTTCAAGTAACACA 480

Qy 1250 ACCGAATCAGACAAAATAATATGGAGCAAGTAGTGAAGAGATGGCTGCTACCGTGC 1309
Db 481 ACCGAATCAGACAAAATAATATGGAGCAAGTAGTGAAGAGATGGCTGCTACCGTGC 540

Qy 1310 TTTGCAAAATGGTGGAACTTACTATATAACCAATGTATCCATAAAGTCGCTTTAGTGAT 1369
Db 541 TTTGCAAAATGGTGGAACTTACTATATAAACAATGTATCCATAAAGTCGCTTTAGTGAT 600

Qy 1370 GGGAGTGAAGAAAGAGTTCTTAATGTCGGAACCTCGTGCCATGAAGGAAACGACAGCCTAT 1429
Db 601 GGGAGTGAAGAAAGAGTTCTTAATGTCGGAACCTCGTGCCATGAAGGAAACGACAGCCTAT 660

Qy 1430 ATGATGACCGA 1440

```



```
|||||
Db 661 ATGATGACCGA 671

RESULT 21
AX110281
LOCUS AX110281 1199 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1014 from Patent WO0123604.
ACCESSION AX110281
VERSION AX110281.1 GI:13926573
KEYWORDS Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
REFERENCE 1 (bases 1 to 1199)
AUTHORS Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 1014 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
FEATURES
source
Location/Qualifiers
1..1199
/organism="Streptococcus pneumoniae"
/strain="StrR-11"
/db_xref="taxon:1313"
BASE COUNT 386 a 286 c 240 g 287 t
ORIGIN
Query Match 29.5%; Score 590; DB 6; Length 1199;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 690; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 761 CAAAACATCTGGGATATTACATACAGAGCAATAGTTGCCCTATCCAGACGATGAA 820
Db 1 CAAAACATCTGGGATATTACAAATACAGAGCAATAGTTGCCCTATCCAGACGATGAA 60
QY 821 TTGCAAGTCGCTTCTACCATTTGTGATGTTCTTAACGGTAAAGTCATGTCGCCAGCTAGGA 880
Db 61 TTGCAAGTCGCTTCTACCATTTGTGATGTTCTTAACGGTAAAGTCATGTCGCCAGCTAGGA 120
QY 881 GCACGCCATCAGTCAAGTAAATGTTCCCTTCGGAAATTAACCAAGCAGTAGAACAACCCG 940
Db 121 GCACGCCATCAGTCAAGTAAATGTTCCCTTCGGAAATTAACCAAGCAGTAGAACAACCCG 180
QY 941 GACTGGGGATCAACTATGAACCGATACAGACTATGCTCCTGCTGGAGTACGGTGC 1000
Db 181 GACTGGGGATCAACTATGAACCGATACAGACTATGCTCCTGCTGGAGTACGGTGC 240
QY 1001 TAGGATTCAACTGCTACTATCGTTACGATGAGCCCTATAACTACCTGGGACAAATACT 1060
Db 241 TAGGATTCAACTGCTACTATCGTTACGATGAGCCCTATAACTACCTGGGACAAATACC 300
QY 1061 CCGTTTATACTGGGATAGGGGCTACTTTGGCAACATCACCTTTGGAATACGCCCTGCAA 1120
Db 301 CCGTTTATACTGGGATAGGGGCTACTTTGGCAACATCACCTTTGGAATACGCCCTGCAA 360
QY 1121 CAATCCGAAACGTCACCGCGGAACTCTAAACAGGTTCGACTCAACCCGCCCAAG 1180
Db 361 CAATCCGAAACGTCACCGCGGAACTCTAAACAGGTTCGACTCAACCCGCCCAAG 420
QY 1181 ACTTTCTAAATGGTCTAGGAATCGACTACCCAAATGATTACACTACCAATGCCATTTC 1240
Db 421 ACTTTCTAAATGGTCTAGGAATCGACTACCCAAATGATTACACTACCAATGCCATTTC 480
QY 1241 AGTAACACACCGAATCAGACAAAATATGGAGCAAGTAGTGAAGAAGTGGCTGCTGCT 1300
Db 481 AGTAACACACCGAATCAGACAAAATATGGAGCAAGTAGTGAAGAAGTGGCTGCTGCT 540
QY 1301 TAGGCTGCCTTTGCAATGGTGAACCTTACTATAACCAATGATATATCAATAAGTTCGTC 1360
|||||
```

```
Db 541 TACGCTGCCCTTGCATAATGCTGGAACCTACTATAAACCAATGATATATCCATAAAGTCGTC 600
QY 1361 TTTAGTGATGGAGTGAAAAAGAGTTCTCTAATGTCGGAACCTGTCCTCAATGAAGGAAACG 1420
|||||
Db 601 TTTAGTGATGGAGTGAAAAAGAGTTCTCTAATGTCGGAACCTGTCCTCAATGAAGGAAACG 660
QY 1421 ACAGCCTATATGATGACCCGACATGATGAAAC 1452
|||||
Db 661 ACAGCCTATATGATGACCCGACATGATGAAAC 692

RESULT 22
AX110284
LOCUS AX110284 1229 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1017 from Patent WO0123604.
ACCESSION AX110284
VERSION AX110284.1 GI:13926576
KEYWORDS Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
REFERENCE 1 (bases 1 to 1229)
AUTHORS Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 1017 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
FEATURES
source
Location/Qualifiers
1..1229
/organism="Streptococcus pneumoniae"
/strain="StrR-14"
/db_xref="taxon:1313"
BASE COUNT 399 a 287 c 240 g 303 t
ORIGIN
Query Match 28.4%; Score 567; DB 6; Length 1229;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 667; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 745 TGTAGACCAAGAGCTCAAAAACATCTGTGGGATATTTTACAATACAGACGAATACGTTGC 804
Db 1 TGTAGACCAAGAGCTCAAAAACATCTGTGGGATATTTTACAATACAGACGAATACGTTGC 60
QY 805 CTATCCAGACGATGAATGCAAGTCGCTTCTACCATTTGTTGATGTTCTTAACGTAAGT 864
Db 61 CTATCCAGACGATGAATGCAAGTCGCTTCTACCATTTGTTGATGTTCTTAACGTAAGT 120
QY 865 CATTTGCCAGCTAGGAGCAGCCCATCAGTCAAGTAATGTTTTCCTTCGGAATTAACCAAGC 924
Db 121 CATTTGCCAGCTAGGAGCAGCCCATCAGTCAAGTAATGTTTTCCTTCGGAATTAACCAAGC 180
QY 925 AGTAGAACAACCCGCGACTGGGGATCAACTATGAACCCGATCAGACACTATGCTCTGTC 984
Db 181 AGTAGAACAACCCGCGACTGGGGATCAACTATGAACCCGATCAGACACTATGCTCTGTC 240
QY 985 CTTGGAGTACGGTGTCTACGATTCACTGCTACTATGTTTCAGATGAGCCCTATATACTA 1044
Db 241 CTTGGAGTACGGTGTCTACGATTCACTGCTACTATGTTTCAGATGAGCCCTATATACTA 300
QY 1045 CCCTGGGACAAATACTCCTGTTTATACTGGGATAGGGGCTACTTTTGGCAACATCACCTT 1104
Db 301 CCCTGGGACAAATACTCCTGTTTATACTGGGATAGGGGCTACTTTTGGCAACATCACCTT 360
QY 1105 GCAATACGCCCTGCACAAATCGCGAAACGTCGCCAGCGGTGGAACTCTAAACAGGTGCG 1164
Db 361 GCAATACGCCCTGCACAAATCGCGAAACGTCGCCAGCGGTGGAACTCTAAACAGGTGCG 420
QY 1165 ACTCAACCCGCCCAAGACTTCTTAATGCTCTAGGAATCGACTACCCAGTATTACCTA 1224
|||||
Db 421 ACTCAACCCGCCCAAGACTTCTTAATGCTCTAGGAATCGACTACCCAGTATTACCTA 480
```



```
QY 1225 CTCAATGCCATTTCAGTAACACAAACCGAATCAGACAAAAAATATGGAGCAAGTAGTGA 1284
|||||
Db 481 CTCAATGCCATTTCAGTAACACAAACCGAATCAGACAAAAAATATGGAGCAAGTAGTGA 540
|||||
QY 1285 AAGATGGCTGCTGCTTACGCTGCCCTTTGCAATGTGGAACCTACTATTAACCAATGTGA 1344
|||||
Db 541 AAGATGGCTGCTGCTTACGCTGCCCTTTGCAATGTGGAACCTACTATTAACCAATGTGA 600
|||||
QY 1345 TATCCATAAAGTCGCTTTAGTAGTGGAGTGGAGTAAAAAGAGTTCTCTAATGTGCGAACTCG 1404
|||||
Db 601 TATCCATAAAGTCGCTTTAGTAGTGGAGTGGAGTAAAAAGAGTTCTCTAATGTGCGAACTCG 660
|||||
QY 1405 TGGCATGAA 1413
|||||
Db 661 TGGCATGAA 669
|||||

RESULT 23
AF046236
LOCUS
DEFINITION Streptococcus pneumoniae strain N94 penicillin-binding protein 1A
ACCESSION AF046236
VERSION AF046236.1 GI:2911417
KEYWORDS
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 930)
Smith,A.M. and Klugman,K.P.
Alterations in pBP 1A essential for high-level penicillin
resistance in Streptococcus pneumoniae
JOURNAL Antimicrob. Agents Chemother. 42 (6), 1329-1333 (1998)
MEDLINE 98287565
AUTHORS Smith,A.M.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1998) Microbiology, SAIMR Hospital, Johannesburg,
Gauteng 2000, South Africa
FEATURES
source
Location/Qualifiers
1..930
/organism="Streptococcus pneumoniae"
/strain="N94"
/db_xref="taxon:1313"
<1..>930
/gene="ponA"
/feature="ponA"
/feature="ponA"
/feature="ponA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1A"
/protein_id="AAC24701.1"
/db_xref="GI:2911418"
/translation="LWDIYNTDEVYVYPDELOVASTIVDSNGKVIQALGARHOSN
VSFGINQAVETNRDWSMTKPTIDYAPALEYGVXDSTATIVHDEPNYPTNTPVYNW
DRGYFENITLQYALQSRNPVAVETLNKVLNRAKTFNLGLIDYPSIHYSNAISSNT
TESDKYGSSEKMAAAYAAAFANGGTYYKPMYTHKVFSDGSEKFEFSNVTGKMKETT
AYMTMKTVLVYIGRGAYLPLPQAGKTGTSYNTDEIEKYIKNTGYVAPDEMVF
GYTRKYAMAVMTGYNRLTPVVGDTVAAKYR"
BASE COUNT 287 a 213 c 201 g 229 t
ORIGIN

Query Match 27.6%; Score 551; DB 1; Length 930;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 651; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 770 CTGTGGGATATTACAAATACAGCAAGTAAGTTCCTATCCAGACCATGAATTCGAAGTC 829
|||||
Db 1 CTGTGGGATATTACAAATACAGCAAGTAAGTTCCTATCCAGACCATGAATTCGAAGTC 60
|||||
```

```
QY 830 GCTTCTACCATTTGTTGATGTTTCTAACGGTAAAGTCAATGCCAGCTAGGAGCACGCCAT 889
|||||
Db 61 GCTTCTACCATTTGTTGATGTTTCTAACGGTAAAGTCAATGCCAGCTAGGAGCACGCCAT 120
|||||
QY 890 CAGTCAAGTAATGTTTCCTTCGGAAATTAACCAAGCAGTAGAACAACACCGAGCTGGGA 949
|||||
Db 121 CAGTCAAGTAATGTTTCCTTCGGAAATTAACCAAGCAGTAGAACAACACCGAGCTGGGA 180
|||||
QY 950 TCAACTATGAAACCGATACAGACTATGCTCCCTGCTTGGAGTACGGTGTCTACGATTCA 1009
|||||
Db 181 TCAACTATGAAACCGATACAGACTATGCTCCCTGCTTGGAGTACGGTGTCTACGATTCA 240
|||||
QY 1010 ACTGCTACTATCGTTCACGATGAGCCCTATACTACCTGGGACAAATACTCTCTGTTAT 1069
|||||
Db 241 ACTGCTACTATCGTTCACGATGAGCCCTATACTACCTGGGACAAATACTCTCTGTTAT 300
|||||
QY 1070 AACTGGGATAGGGCTACTTTGGCAACATCACCTTGCATACGCCCTGCAACATCGGA 1129
|||||
Db 301 AACTGGGATAGGGCTACTTTGGCAACATCACCTTGCATACGCCCTGCAACATCGGA 360
|||||
QY 1130 AACGTCGCCAGCCGTGGAACCTCTAAACAGGTGCGACTCAACCGCCCAAGACTTTTCTA 1189
|||||
Db 361 AACGTCGCCAGCCGTGGAACCTCTAAACAGGTGCGACTCAACCGCCCAAGACTTTTCTA 420
|||||
QY 1190 AATGCTCTAGGAATCGACTACCCCAAGTATTCACTACTCAAAATGCCATTTCAAGTAACACA 1249
|||||
Db 421 AATGCTCTAGGAATCGACTACCCCAAGTATTCACTACTCAAAATGCCATTTCAAGTAACACA 480
|||||
QY 1250 ACCGAATCAGACAAAAAATATGGAGCAAGTAGTGAAGAGTGGCTGCTGTTAGCTGCC 1309
|||||
Db 481 ACCGAATCAGACAAAAAATATGGAGCAAGTAGTGAAGAGTGGCTGCTGTTAGCTGCC 540
|||||
QY 1310 TTTCGAAATGGTGGAACCTACTATAAACCAATGTATATCCATAAAGTCGTTTAGTGAT 1369
|||||
Db 541 TTTCGAAATGGTGGAACCTACTATAAACCAATGTATATCCATAAAGTCGTTTAGTGAT 600
|||||
QY 1370 GGGAGTGAAGAGAGTTCTCTAATGTGCGAATCTGTCGCCATGAAGGAACGAC 1422
|||||
Db 601 GGGAGTGAAGAGAGTTCTCTAATGTGCGAATCTGTCGCCATGAAGGAACGAC 653
|||||

RESULT 24
AF046234
LOCUS
DEFINITION Streptococcus pneumoniae strain 56739 penicillin-binding protein 1A
ACCESSION AF046234
VERSION AF046234.1 GI:2911413
KEYWORDS
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 930)
Smith,A.M. and Klugman,K.P.
Alterations in pBP 1A essential for high-level penicillin
resistance in Streptococcus pneumoniae
JOURNAL Antimicrob. Agents Chemother. 42 (6), 1329-1333 (1998)
MEDLINE 98287565
AUTHORS Smith,A.M.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1998) Microbiology, SAIMR Hospital, Johannesburg,
Gauteng 2000, South Africa
FEATURES
source
Location/Qualifiers
1..930
/organism="Streptococcus pneumoniae"
/strain="56739"
/db_xref="taxon:1313"
<1..>930
/feature="ponA"
/feature="ponA"
/feature="ponA"
/feature="ponA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1A"
/protein_id="AAC24701.1"
/db_xref="GI:2911418"
/translation="LWDIYNTDEVYVYPDELOVASTIVDSNGKVIQALGARHOSN
VSFGINQAVETNRDWSMTKPTIDYAPALEYGVXDSTATIVHDEPNYPTNTPVYNW
DRGYFENITLQYALQSRNPVAVETLNKVLNRAKTFNLGLIDYPSIHYSNAISSNT
TESDKYGSSEKMAAAYAAAFANGGTYYKPMYTHKVFSDGSEKFEFSNVTGKMKETT
AYMTMKTVLVYIGRGAYLPLPQAGKTGTSYNTDEIEKYIKNTGYVAPDEMVF
GYTRKYAMAVMTGYNRLTPVVGDTVAAKYR"
BASE COUNT 287 a 213 c 201 g 229 t
ORIGIN
```



```
/gene="ponA"  
/note="PBP 1A"  
/codon_start=1  
/transl_table=11  
/product="penicillin-binding protein 1A"  
/protein_id="AAC24699.1"  
/db_xref="GI:2911414"  
/translation="LWDIYNTDEYVAYPDELOQVASTIVDSNGKVIAGOLGARHOSN  
VSFGINQAVETNRDGMSTKPTIDYAPALEYGVYDSTATIVHDEPNYVPGTNPVYNW  
DRGYEGNTLOVALQOSRNPVAVETLNKVLNRKTFELGLGIDYPSIHSYNAISNT  
TESDKKYGASSEKMAAAYAAPANGCTYKPMYIHKVFSDESGSEKESDAGTRAMKETT  
AYMTEMKTVLYIGRGAYLPWLPQAGKTGTSNVTDEIEKYIKNTGYVAPDEMVF  
GYTRYAMAVWTGYSNRLTPLVNGLTVAAKVYR"  
BASE COUNT      288 a      216 c      196 g      230 t  
ORIGIN
```

```
Query Match      23.1%; Score 461; DB 1; Length 930;  
Best Local Similarity 99.5%; Pred. No. 2.3e-248;  
Matches 611; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 770 CTGTGGGATATTACAAATACAGAGTAATACGTTGCCTATCCAGACGATGAATTCGAAGTC 829  
DB 1 CTGTGGGATATTACAAATACAGAGTAATACGTTGCCTATCCAGACGATGAATTCGAAGTC 60  
  
QY 830 GCTTCTACCATTGTTGATGTTCTTAACGGTAAAGTCAATTCGCCAGCTAGGAGCAGCCCAT 889  
DB 61 GCTTCTACCATTGTTGATGTTCTTAACGGTAAAGTCAATTCGCCAGCTAGGAGCAGCCCAT 120  
  
QY 890 CAGTCAAGTAAGTTCCTTCGGAATTAACCAAGCAGTAGAACAACACCGCAGCTGGGA 949  
DB 121 CAGTCAAGTAAGTTCCTTCGGAATTAACCAAGCAGTAGAACAACACCGCAGCTGGGA 180  
  
QY 950 TCAACTATGAAACCGGATCACAGACTAGTCTCTGCTTGGAGTACGGTGTCTACGATTCA 1009  
DB 181 TCAACTATGAAACCGGATCACAGACTAGTCTCTGCTTGGAGTACGGTGTCTACGATTCA 240  
  
QY 1010 ACTGCTACTATGTTCCAGATGAGCCCTTAATACTACCTCGGGACAATACTCTGTTTAT 1069  
DB 241 ACTGCTACTATGTTCCAGATGAGCCCTTAATACTACCTCGGGACAATACTCTGTTTAT 300  
  
QY 1070 AACTGGGATAGGGCTACTTTGGCAACATCACCTTGCATACGCCCTGCAACAATCGGA 1129  
DB 301 AACTGGGATAGGGCTACTTTGGCAACATCACCTTGCATACGCCCTGCAACAATCGGA 360  
  
QY 1130 AAGTCCCGAGCGGTGAAACTCTAAACAGGTCGGACTCAACCGCGCAAGACTTTCCCTA 1189  
DB 361 AAGTCCCGAGCGGTGAAACTCTAAACAGGTCGGACTCAACCGCGCAAGACTTTCCCTA 420  
  
QY 1190 AATGGTCTAGGAATGACTTACCCAGTATTCATCTACTCAAAATGCCATTTCAAGTAACACA 1249  
DB 421 AATGGTCTAGGAATGACTTACCCAGTATTCATCTACTCAAAATGCCATTTCAAGTAACACA 480  
  
QY 1250 ACCGAATCAGACAAAATAATGGAGCAAGTAGTGAAAAGATGGCTGCTTACGCTGCC 1309  
DB 481 ACCGAATCAGACAAAATAATGGAGCAAGTAGTGAAAAGATGGCTGCTTACGCTGCC 540  
  
QY 1310 TTTGCAAAATGGTGGAACTTACTATATAACCAATGTATATCCATAAGTCTGCTTTAGTGAT 1369  
DB 541 TTTGCAAAATGGTGGAACTTACTATATAACCAATGTATATCCATAAGTCTGCTTTAGTGAT 600  
  
QY 1370 GGGAGTGA AAAAGA 1383  
DB 601 GGGAGTGA AAAAGA 614  
  
RESULT 25  
AF210747  
LOCUS  
DEFINITION  
2160 bp DNA linear BCT 17-SEP-2001  
Streptococcus pneumoniae isolate SP 1465 penicillin-binding protein  
1A (pbpla) gene, complete cds.  
ACCESSION  
AF210747  
VERSION  
AF210747.1 GI:6563340
```

```
KEYWORDS  
SOURCE Streptococcus pneumoniae.  
ORGANISM Streptococcus pneumoniae  
Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
Streptococcus.  
REFERENCE 1 (bases 1 to 2160)  
AUTHORS Ferroni, A. and Berche, P.  
TITLE Alterations to penicillin-binding proteins 1A, 2B and 2X amongst  
penicillin-resistant clinical isolates of Streptococcus pneumoniae  
serotype 23F from the nasopharyngeal flora of children  
J. Med. Microbiol. 50 (9), 828-832 (2001)  
JOURNAL 21432820  
MEDLINE  
PUBMED 11549185  
REFERENCE 2 (bases 1 to 2160)  
AUTHORS Ferroni, A. and Berche, P.  
TITLE Direct Submission  
JOURNAL Submitted (03-DEC-1999) Microbiology, Hospital  
Necker-Enfants-Malades, 149 rue de Sevres, Paris 75015, France  
FEATURES  
Location/Qualifiers  
1..2160  
/organism="Streptococcus pneumoniae"  
/serotype="23F"  
/isolate="SP 1465"  
/db_xref="taxon:1313"  
1..2160  
/gene="pbpla"  
1..2160  
/gene="pbpla"  
/codon_start=1  
/transl_table=11  
/product="penicillin-binding protein 1A"  
/protein_id="AAF17257.1"  
/db_xref="GI:6563341"  
/translation="MNKPTILRLIKVLSIFLSLVIAIVLGGVFFVYVSKAPLSLSE  
SKLIVATSSKIYDNKQLIADLGSERRVNAQADIPDILVKAIVSTEDHRRFDHREGID  
TIRLIGALRLNQSNSLOGSGSTLTQOLIKLTYESTSDQTIISRAQEAWLAIQLEQK  
ATKQELIYIINKYMSNGMTAAQNYIGRLRELSPQLALLAGMPQAPNQIDP  
YSHPEAALDRRLVLSKMGQGYISAEQYKAINTPTDGLQSLKSVNSYPAYMONYL  
KEVIDQEQTGYNLITGMQYITNVQEAQKLMWDIYNSDQYVSPDDDLQVASTVY  
VNSDKPIAQLGARHQASNVSGFTNQAVETNRDGMSTKPTIDYAPALEYGVYDSTAT  
MYNDIPYNYPGTSTPVYNWDRYFGNITLOVALQOSRNPVAVETLNKVLNRKTFEL  
GLGIDYPODMHYSNAJSSNTESKQYGAASSEKMAAFAAFANGLYHKPMYINKLVFS  
DSESEFSDAGTRAMKETTAYMTEMKTVLYSGTGRNAYLAWLFOAGKTGTSNTTDE  
ELENHKTQSQVAPDELFAGYTRKYSMAVYGTYSNRLTPLVNGLTVAAKVYRSMWY  
LSEGSNEDWNIPEGLYRNGEYFVKNGARSTWSSPAQPPQPPSTESSSSSSSSSTSSQSS  
STTPSTNNSTTTNNNTQQSTTPDQONQNPQAP"  
BASE COUNT      692 a      505 c      405 g      558 t  
ORIGIN
```

```
Query Match      21.8%; Score 436; DB 1; Length 2160;  
Best Local Similarity 99.6%; Pred. No. 3e-234;  
Matches 536; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1462 TTATGGAACGTGACGAAATGCCTATCTTGGTCCCTCAGGCTGTTAAACACGGAAC 1521  
DB 1620 TTATGGAACGTGACGAAATGCCTATCTTGGTCCCTCAGGCTGTTAAACACGGAAC 1679  
  
QY 1522 CTCTAACTATACAGACGAGGAATTTGAAACACACATCAAGACCTCTCAATTTTACGACC 1581  
DB 1680 CTCTAACTATACAGACGAGGAATTTGAAACACACATCAAGACCTCTCAATTTTACGACC 1739  
  
QY 1582 TGATGAACATTTTGGCTGCTATACGGGTAAATATTTCAATGGCTGTATGGACAGGCTATTC 1641  
DB 1740 TGATGAACATTTTGGCTGCTATACGGGTAAATATTTCAATGGCTGTATGGACAGGCTATTC 1799  
  
QY 1642 TAACCGTCTGACACCACTTGTAGGCAATGGCCTTAGGCTGCTGCCAAAGTTTACCGCTC 1701  
DB 1800 TAACCGTCTGACACCACTTGTAGGCAATGGCCTTAGGCTGCTGCCAAAGTTTACCGCTC 1859  
  
QY 1702 TATGATGACCTACCTCTCTCAAGGAAGCAATCCAGAAAGATTGGAATATATACGAGGGGCT 1761  
DB 1860 TATGATGACCTACCTCTCTCAAGGAAGCAATCCAGAGGATTGGAATATATACGAGGGGCT 1919
```



```
QY 1762 CTACAGAAATGGAGATTTCGTATTTAAAAATGGTGTCTTCTTACGTGGAATCACTGTC 1821
|||||
Db 1920 CTACAGAAATGGAGATTTCGTATTTAAAAATGGTGTCTTCTTACGTGGAATCACTGTC 1799
|||||
QY 1822 TCACAACAACCCCCATCAACTGAAAGTTCAAGTCATCATCATAGATAGTTCAACTTCACA 1881
|||||
Db 1980 TCACAACAACCCCCATCAACTGAAAGTTCAAGTCATCATCATAGATAGTTCAACTTCACA 2039
|||||
QY 1882 GTCTAGCTCAACCACTCCCAAGCACAATAATAGTAGGACTACCAATCCTTAACAATAATAC 1941
|||||
Db 2040 GTCTAGCTCAACCACTCCCAAGCACAATAATAGTAGGACTACCAATCCTTAACAATAATAC 2099
|||||
QY 1942 GCAACAATCAATACAAACCCCTGATCAACAAAATCAGAATCCTCAACACGACACAACA 1999
|||||
Db 2100 GCAACAATCAATACAAACCCCTGATCAACAAAATCAGAATCCTCAACACGACACAACA 2157

RESULT 26
SPNEU1914/c 30457 bp DNA linear HTG 11-JUL-2001
LOCUS Streptococcus pneumoniae clone G54, *** SEQUENCING IN PROGRESS ***,
DEFINITION in ordered pieces.
ACCESSION AL449936
VERSION AL449936.1 GI:11545161
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 30457)
Dopazo,J., Mendoza,A., Herrero,J., Caldara,F., Polissi,A.,
Friedli,L., Guerrier,M., Grand-Schenk,E., Gandin,C., de
Francesco,M., Polissi,A., Buell,G., Feger,G., Garcia,E., Peitsch,M.
and Garcia-Bustos,J.F.
Annotated draft genomic sequence from a Streptococcus pneumoniae
type 19F clinical isolate
Microb. Drug Resist. 7 (2), 99-125 (2001)
21335329
2 (bases 1 to 30457)
Dopazo,J., Mendoza,A., Herrero,J., Caldara,F., Polissi,A.,
Humbert,Y., Friedli,L., Guerrier,M., Grand-Schenk,E., Gandin,C., de
Francesco,M., Buell,G., Feger,G., Garcia,E., Peitsch,M. and
Garcia-Bustos,J.F.
Direct Submission
Submitted (31-OCT-2000) Research Department, Glaxo Wellcome, S.A.,
Severo Ochoa 2, 28760 Tres Cantos, SPAIN
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1. 30457
/organism="Streptococcus pneumoniae"
/serotype="19F"
/db_xref="taxon:1313"
/clone="G54"

BASE COUNT 9377 a 5459 c 6906 g 8705 t 10 others
ORIGIN

Query Match 21.8%; Score 436; DB 2; Length 30457;
Best Local Similarity 99.6%; Pred. No. 2.9e-234;
Matches 536; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1462 TTATGGAATGCACGAAATGCTATCTGCTTGGCTCCCTCAGGCTGTTAAACAGGAAC 1521
|||||
Db 8148 TTATGGAATGCACGAAATGCTATCTGCTTGGCTCCCTCAGGCTGTTAAACAGGAAC 8089
|||||
QY 1522 CTCCTACTACAGAGGAGAAATGAAACACACATCAAGACTCTCAATTTGTAGCACC 1581
|||||
Db 8088 CTCCTACTACAGAGGAGAAATGAAACACACATCAAGACTCTCAATTTGTAGCACC 8029
|||||
```

```
QY 1582 TGATGAACATTTTGTCTGCTATACCGCTAAATATTCAATGGTGTATGACAGGCTATTTC 1641
|||||
Db 8028 TGATGAACATTTTGTCTGCTATACCGCTAAATATTCAATGGTGTATGACAGGCTATTTC 7969
|||||
QY 1642 TTACCGTCTGACACCACTTGTAGGCAATGGCTTACGGTCGTCGCCAAAGTTTACCGCTC 1701
|||||
Db 7968 TTACCGTCTGACACCACTTGTAGGCAATGGCTTACGGTCGTCGCCAAAGTTTACCGCTC 7909
|||||
QY 1702 TATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAAAGATTGGAATATACCAGAGGGCT 1761
|||||
Db 7908 TATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAAAGATTGGAATATACCAGAGGGCT 7849
|||||
QY 1762 CTACAGAAATGGAGATTTCGTATTTAAAAATGGTGTCTTCTTACGTGGAATCACTGTC 1821
|||||
Db 7848 CTACAGAAATGGAGATTTCGTATTTAAAAATGGTGTCTTCTTACGTGGAATCACTGTC 7789
|||||
QY 1822 TCACAACAACCCCCATCAACTGAAAGTTCAAGTCATCATCATAGATAGTTCAACTTCACA 1881
|||||
Db 7788 TCACAACAACCCCCATCAACTGAAAGTTCAAGTCATCATCATAGATAGTTCAACTTCACA 7729
|||||
QY 1882 GTCTAGCTCAACCACTCCCAAGCACAATAATAGTAGGACTACCAATCCTTAACAATAATAC 1941
|||||
Db 7728 GTCTAGCTCAACCACTCCCAAGCACAATAATAGTAGGACTACCAATCCTTAACAATAATAC 7669
|||||
QY 1942 GCAACAATCAATACAAACCCCTGATCAACAAAATCAGAATCCTCAACACGACACAACA 1999
|||||
Db 7668 GCAACAATCAATACAAACCCCTGATCAACAAAATCAGAATCCTCAACACGACACAACA 7611
|||||

RESULT 27
AF210752 2160 bp DNA linear BCT 17-SEP-2001
LOCUS Streptococcus pneumoniae isolate BM 4200 penicillin-binding protein
DEFINITION 1A (pbpla) gene, complete cds.
ACCESSION AF210752
VERSION AF210752.1 GI:6563350
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 2160)
Ferroni,A. and Berche,P.
Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
penicillin-resistant clinical isolates of Streptococcus pneumoniae
serotype 23F from the nasopharyngeal flora of children
J. Med. Microbiol. 50 (9), 828-832 (2001)
21432820
JOURNAL MEDLINE
PUBMED 11549185
REFERENCE 2 (bases 1 to 2160)
FERRONI,A. and BERCHE,P.
AUTHORS Direct Submission
TITLE Submitted (03-DEC-1999) Microbiology, Hospital
Necker-Enfants-Malades, 149 rue de Sevres, Paris 75015, France
JOURNAL Location/Qualifiers
FEATURES
source
1. 2160
/organism="Streptococcus pneumoniae"
/serotype="23F"
/isolate="BM 4200"
/db_xref="taxon:1313"
/note="strain isolated in 1978 in Paris"
1. 2160
/gene="pbpla"
1. 2160
/gene="pbpla"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1A"
/protein_id="AAF17262.1"
/db_xref="GI:6563351"
/translaton="MNKPTILRLKYLISFLSILVIAALVGGGVFFVYVSKAPLSLSE
SKLVATTSKLYDNKNQLIADLGSERRVNAQANDIPTDLVKAIVSIEDHRFFDHRGID
TIRILGAFRLNLSNSLOGGSTLTQOLIKLTYFSTSTSDQITISRAQEAQLAIQEQ
```


ATKQILYYINKVMSNGNTGMQTAQNYGKDLNNLSLPOLALLAGMPOAPNOYDP
YSHPAQRDRNLRLISEMKNQYISAEQYERAVNPTITDGLQSLKSASNYPAYMDNYL
KEYINQVEEETGYNLLTGMVYTDQEAQKHLWDIYNTDEYVAYPDELOVASTIV
DVSNGKVIQAOLCARHOSNSVFGINOAVENTNRDWSMTKPTIDYAPALEYGVYDSTAT
IVHDEPYNYPGTDPVYNDRGRYFGNITLOVALQOSRNPVAVETLNKYLCLNRAKTEL
GLGIDYPSLHYSNALSNTTSDKKYGASSEKMAAAYAFANGYTYYPWYHKKYVES
DGSEKFSNVGRANKETATMTDMKTIVLYGIGRGAYLPLWLPQAKTGYSNTYDE
LIEGSKNEDWIPEDLYRNGEFVFNKARSTWSSPAQPPSTESSSSSSDSTSQSN
STTSPSTNNSTTTNNNTTQSNNTTPOQNPQPAQ*
BASE COUNT 698 a 520 c 416 g 526 t
ORIGIN

Query Match 21.6%; Score 432; DB 1; Length 2160;
Best Local Similarity 98.9%; Pred. No. 5.4e-232;
Matches 1032; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATAAATAAATCAACTCATTCGTGACTTGGGTTCGAAAGCGCGGTCAA 60
|||||
DB 159 TAAATCTACGACATAAATAAATCAACTCATTCGTGACTTGGGTTCGAAAGCGCGGTCAA 218
|||||
QY 61 TGCCCAAGCTAATGATATTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
|||||
DB 219 TGCCCAAGCTAATGATATTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
|||||
QY 121 TCGCTTCTTCGACCAGGGGATGATACCATCCGTATCCCTGGGAGCTTCTTCGCGCAA 180
|||||
DB 279 TCGCTTCTTCGACCAGGGGATGATACCATCCGTATCCCTGGGAGCTTCTTCGCGCAA 338
|||||
QY 181 TCTGCAAGCAATTCCTCCAGGTGGATCACTCFACCCCAACAGTTGATTAAGTTGAC 240
|||||
DB 339 TCTGCAAGCAATTCCTCCAGGTGGATCACTCFACCCCAACAGTTGATTAAGTTGAC 398
|||||
QY 241 TTACTTTTCAACTTCGACTTCGACAGACTATTCTCTCTAGGCTCAGSAGCTTGGTT 300
|||||
DB 399 TTACTTTTCAACTTCGACTTCGACAGACTATTCTCTCTAGGCTCAGSAGCTTGGTT 458
|||||
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
|||||
DB 459 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 518
|||||
QY 361 GGTCTACATGCTAATGGGAATATGGAATGGAGACAGAGCTCAAACTACTATGGTAA 420
|||||
DB 519 GGTCTACATGCTAATGGGAATATGGAATGGAGACAGAGCTCAAACTACTATGGTAA 578
|||||
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGTGCTGGCTGGAATCGCTCAGGCACC 480
|||||
DB 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGTGCTGGCTGGAATCGCTCAGGCACC 638
|||||
QY 481 AAACCAATATGACCCCTATTTCACATCCAGAAGCAGCCCAAGACCCGCGAAACTTGGCTTT 540
|||||
DB 639 AAACCAATATGACCCCTATTTCACATCCAGAAGCAGCCCAAGACCCGCGAAACTTGGCTTT 698
|||||
QY 541 ATCTGAATGAAATFCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
|||||
DB 699 ATCTGAATGAAATFCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 758
|||||
QY 601 ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAAGTAATTACCCCTGCTTACATGGA 660
|||||
DB 759 ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAAGTAATTACCCCTGCTTACATGGA 818
|||||
QY 661 TAAATACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAAGAAACAGGCTATAAACCCTACTCAC 720
|||||
DB 819 TAAATACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAAGAAACAGGCTATAAACCCTACTCAC 878
|||||
QY 721 AACTGGGATGGATGCTACACAAATGTAGACCAAGAGCTCAAAAACATCTCTGGGATAT 780
|||||
DB 879 AACTGGAATGGGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTCTGGGATAT 938
|||||
QY 781 TTACAATACAGCAAGTAATCGTTCCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 840
|||||
DB 939 TTACAATACAGCAAGTAATCGTTCCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 998
|||||

QY 841 TGTGATGTTTCTAACGGTAAAGTCATTGCCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
|||||
DB 999 TGTGATGTTTCTAACGGTAAAGTCATTGCCCAGCTAGGAGCAGCCATCAGTCAAGTAA 1058
|||||
QY 901 TGTTTCTTCGGAATTAACCAAGCAGTAGAACAACCCGCGACTGGGGATCAACTATGAA 960
|||||
DB 1059 TGTTTCTTCGGAATTAACCAAGCAGTAGAACAACCCGCGACTGGGGATCAACTATGAA 1118
|||||
QY 961 ACCGATCACAGACTATGCTCCTGCTGGAGTACGGTGTCTAGCATTCAACTGCTACTAT 1020
|||||
DB 1119 ACCGATCACAGACTATGCTCCTGCTGGAGTACGGTGTCTAGCATTCAACTGCTACTAT 1178
|||||
QY 1021 CGTTCAGATGAGCCCTATATAACTA 1044
|||||
DB 1179 CGTTCAGATGAGCCCTATATAACTA 1202
|||||
RESULT 28
AX111314
LOCUS
DEFINITION Sequence 2047 from Patent WO0123604.
ACCESSION AX111314
VERSION AX111314.1 GI:13927606
KEYWORDS
SOURCE
ORGANISM Streptococcus pneumoniae.
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 2047 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
FEATURES
source
1..2160
/organism="Streptococcus pneumoniae"
/strain="CS109"
/db_xref="taxon:1313"
BASE COUNT 701 a 494 c 395 g 570 t
ORIGIN

Query Match 19.8%; Score 396; DB 6; Length 2160;
Best Local Similarity 99.1%; Pred. No. 1.2e-211;
Matches 696; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 TAAATCTACGACATAAATAAATCAACTCATTCGTGACTTGGGTTCGAAAGCGCGGTCAA 60
DB 159 TAAATCTACGACATAAATAAATCAACTCATTCGTGACTTGGGTTCGAAAGCGCGGTCAA 218
|||||
QY 61 TGCCCAAGCTAATGATATTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
DB 219 TGCCCAAGCTAATGATATTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
|||||
QY 121 TCGCTTCTTCGACCAGGGGATGATACCATCCGTATCCCTGGGAGCTTCTTCGCGCAA 180
DB 279 TCGCTTCTTCGACCAGGGGATGATACCATCCGTATCCCTGGGAGCTTCTTCGCGCAA 338
|||||
QY 181 TCTGCAAGCAATTCCTCCAGGTGGATCACTCFACCCCAACAGTTGATTAAGTTGAC 240
DB 339 TCTGCAAGCAATTCCTCCAGGTGGATCACTCFACCCCAACAGTTGATTAAGTTGAC 398
|||||
QY 241 TTACTTTTCAACTTCGACTTCGACAGACTATTCTCTGTAAGGCTCAGGAAGCTTGGTT 300
DB 399 TTACTTTTCAACTTCGACTTCGACAGACTATTCTCTGTAAGGCTCAGGAAGCTTGGTT 458
|||||
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
DB 459 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 518
|||||


```
QY 361 GGTCTACATGCTTAATGGGAACCTATGAATGCAGACAGCAGCTCAAACTACTACTATGGTAA 420
Db 519 GGTCTACATGCTTAATGGGAACCTATGAATGCAGACAGCAGCTCAAACTACTACTATGGTAA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGTCTGGCTGGGATGCTCAGGCACC 480
Db 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGTCTGGCTGGGATGCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTACATCCAGAACGAGCCCAAGACCGCGGAAACTTGTCTTT 540
Db 639 AAACCAATATGACCCCTATTACATCCAGAACGAGCCCAAGACCGCGGAAACTTGTCTTT 698
QY 541 ATCTGAAATGAAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGCTCAATAC 600
Db 699 ATCTGAAATGAAATAATCAAGGCTTACATCTCTGCTGAACAGTATGAGAAAGCAGCTCAATAC 758
QY 601 ACCAATTTACTGATGGACTACAAAGTCTCAATCAGCAAGTAAATACCTGCTTACATGGA 660
Db 759 ACCAATTTACTGATGGACTACAAAGTCTCAATCAGCAAGTAAATACCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGAAGTCTCAATCAATCAAGTTGAAGAAGAAAC 702
Db 819 TAATTACCTCAAGGAAGTCTCAATCAATCAAGTTGAAGAAGAAAC 860

RESULT 29
SPCS1091A SPCS1091A 2172 bp DNA linear BCT 08-SEP-1995
LOCUS S.pneumoniae CS109 pbpla gene for penicillin-binding protein la.
DEFINITION 249094
ACCESSION 249094
VERSION 249094.1 GI:984228
KEYWORDS PBP 1A gene; penicillin-binding protein la.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 2172)
Coffey,T.J., Daniels,M., McDougal,L.K., Dowson,C.G., Tenover,F.C.
and Spratt,B.G.
Genetic analysis of clinical isolates of Streptococcus pneumoniae
with high-level resistance to expanded-spectrum cephalosporins
Antimicrob. Agents Chemother. 39 (6), 1306-1313 (1995)
96012191
2 (bases 1 to 2172)
Direct Submission
Spratt,B.G.
Submitted (21-APR-1995) Spratt B. G., University of Sussex, School
of Biology, Falmer, Brighton, East Sussex, UK, BN1 9QG
location/Qualifiers
1. .2172
/organism="Streptococcus pneumoniae"
/strain="CS109 (cephalosporin-resistant clinical isolate)"
/db_xref="taxon:1313"
13. .2172
/gene="pbpla"
/gene="pbpla"
/function="peptidoglycan transglycosylase-transpeptidase"
/citation=[1]
/codon_start=1
/transl_table=11
/evidence=experimental
/product="penicillin-binding protein la"
/protein_id="CAA88917.1"
/db_xref="GI:984229"
/db_xref="SPTREMBL:Q57114"
/translation="MNKPTILRLIKYLSIFLSLVIAIVLGGGVFFVYVSKAPLSLSE
SKLVATTSKIYDNKNLIADISERRVNAQADPTDLVKAIVSIEDHRRFDFHRGID
TTRILGAFILNQSLOGSGSTLQOLIKLTVFSTSDOTSRAQAEAWLAIOLEQK
ATFOELITYINKVYMGNGYGMQTAQNYGKDLNLSLPOLALLAGMPOAPNOYDP
YSHPEAODRNLVLSEMKNOGYISAEQYKAVNPITDGLQSLKASNPAYMNYL
KEVINQVEETGYNLTITGMQVYVITNDQEAQKHLWDIYNSDQIVSYFDDLDQVASTV
DVSNGKVIQAQGARHQASNVSGFTNQAVETNRDWSMKRPIITDYAPALEYGYVDSTAS
```

```
IVHDVPYNYPGTDTPLYNWHDVYFGNITIQYALQOOSRNVYAVETLKNKVLGRKTELN
GLGIDYPSMHYANAISSTNTESNKRYGASSERKMAAAYAAFAAGGIYHKPMYINKVIFS
DCSEKFEVDAGTRAMKETTYNMTEMKTVLTVGTGRGAYLWLPQAGKTGTSNYTDE
ELEKYIKNTGYVAPDEMFGVYTRKYAMAYMTGYSNRLTPIIGDGLVACKVVYSMLTY
LSEDDQPGDWTMPDGLYRNGEVEVFNKARSTWSSPAPQPPPTSESSSSSDSSTSUSN
STPSTNNSTTTNNNTTTPDQONQNPQAPQ"
1120. .1122
/misc_feature
/gene="pbpla"
/note="active site"
BASE COUNT 707 a 495 c 398 g 572 t
ORIGIN
Query Match 19.8%; Score 396; DB 1; Length 2172;
Best Local Similarity 99.1%; Pred. No. 1.2e-211;
Matches 696; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 TAAATCTAGGCAATAAATAATCAATCAATGCTGCTGAGCTTGGGTCTGGAAGCGCGGTCAA 60
Db 171 TAAATCTAGGCAATAAATAATCAATCAATGCTGCTGAGCTTGGGTCTGGAAGCGCGGTCAA 230
QY 61 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 231 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 290
QY 121 TCGCTTCTTCGACCACAGGGGGATTGATACCATCCGCTATCCTGGAGCTTCTTTCGCGCAA 180
Db 291 TCGCTTCTTCGACCACAGGGGGATTGATACCATCCGCTATCCTGGAGCTTCTTTCGCGCAA 350
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCAACTCTCACCAACAGTTGATTAAGTTGAC 240
Db 351 TCTGCAAGCAATTCCTCCCAAGGTGATCAACTCTCACCAACAGTTGATTAAGTTGAC 410
QY 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 411 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 470
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
Db 471 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 530
QY 361 GGTCTACATGCTTAATGGGAACCTATGGAATGCAGACAGCAGCTCAAACTACTATGGTAA 420
Db 531 GGTCTACATGCTTAATGGGAACCTATGGAATGCAGACAGCAGCTCAAACTACTATGGTAA 590
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGTGGCTGGAATGCTCAGGCACC 480
Db 591 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGTGGCTGGAATGCTCAGGCACC 650
QY 481 AAACCAATATGACCCCTATTACATCCAGAACGAGCCCAAGACCGCGGAAACTTGTCTTT 540
Db 651 AAACCAATATGACCCCTATTACATCCAGAACGAGCCCAAGACCGCGGAAACTTGTCTTT 710
QY 541 ATCTGAAATGAAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGCTCAATAC 600
Db 711 ATCTGAAATGAAATAATCAAGGCTTACATCTCTGCTGAACAGTATGAGAAAGCAGCTCAATAC 770
QY 601 ACCAATTTACTGATGGACTACAAAGTCTCAATCAGCAAGTAAATACCTGCTTACATGGA 660
Db 771 ACCAATTTACTGATGGACTACAAAGTCTCAATCAGCAAGTAAATACCTGCTTACATGGA 830
QY 661 TAATTACCTCAAGGAAGTCTCAATCAATCAAGTTGAAGAAGAAAC 702
Db 831 TAATTACCTCAAGGAAGTCTCAATCAATCAAGTTGAAGAAGAAAC 872

RESULT 30
SPCS1111A SPCS1111A 2172 bp DNA linear BCT 05-JUN-1997
LOCUS S.pneumoniae CS111 pbpla gene for penicillin-binding protein la.
DEFINITION 249095
ACCESSION 249095
VERSION 249095.1 GI:984230
KEYWORDS pbpla; penicillin-binding protein la.
```



```
SOURCE      Streptococcus pneumoniae.
ORGANISM     Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
REFERENCE    1 (bases 1 to 2172)
AUTHORS      Spratt,B.G.
TITLE        Direct Submission
JOURNAL      Submitted (21-APR-1995) Spratt B. G., University of Sussex, School
MEDLINE      of Biology, Falmer, Brighton, East Sussex, UK, BN1 9QG
REFERENCE    2 (bases 1 to 2172)
AUTHORS      Coffey,T.J., Daniels,M., McDougal,L.K., Dowson,C.G., Tenover,F.C.
TITLE        and Spratt,B.G.
JOURNAL      Genetic analysis of clinical isolates of Streptococcus pneumoniae
MEDLINE      with high-level resistance to expanded-spectrum cephalosporins
FEATURES     Antimicrob. Agents Chemother. 39 (6), 1306-1313 (1995)
96012191
Location/Qualifiers
  source      1..2172
              /organism="Streptococcus pneumoniae"
              /strain="CS111 (cephalosporin-resistant clinical isolate)"
              /db_xref="taxon:1313"
  gene        13..2172
              /gene="pbpla"
  CDS         13..2172
              /gene="pbpla"
              /standard_name="pbpla"
              /function="peptidoglycan transglycosylase-transpeptidase"
              /codon_start=1
              /transl_table=11
              /evidence=experimental
              /product="penicillin-binding protein 1a"
              /protein_id="CAA88918.1"
              /db_xref="GI:984231"
              /db_xref="SPTREMBL:Q57114"
              /translation="MNKPTILRLKYLISFLSLVIAIVLGGVFFVYVSKAPLSLSE
SKLVATTSKLYDNKNQLIADLGSERRVNAQNDIPTDLVRAIVSIEDHRRFDHGRGID
TIRLGAFLRNLNSLQSGSTLTQQLIKLTYFSTSDQTSIRKRAQAWLAIAQLQEQ
ATKQELIYYLNKVMYSGNGYGMOTAAQNYGKOLNLSLPOLALQAQPOAQNQYDP
YSHPEADQNRNLYLSEMKNGYISAEQYKAVNPPIYDGLQSLKASNYPAYMDNYL
KEVINQVEETGYNLLTTGMDVYTVNDOEAKHLWDIYNSDYQYVYVDDQLQVASTVS
DVSKVYTAOLGARHQASNVSEFTNQVAVETNRDMSMKPIYDVAPELEYGVYDSTAS
IVHDVPYNYPGTDPLYNWDHYFNGNIYIYALQSRNVAVETLNKVLGRKRAKFTLN
GLGIDPSMHYANALISNTSNKYGASSERMAAYAFANGGIYHHPKMTINKIVFS
DGSEKFSADGATRAKMETAYMTMMKTVLTGYTGRGAYILPWLPAQKGTTSNTDE
BIERYIKNTGVADPEMVEGYTRKYAMAVWTGYSNRLTPIIGDGLVAGKTVYRSMITY
LSEDDQPDGMDGLYRNGEYFKNGARSTWSSPAQOPSTSESSSSSDSSTQSQN
STTPSTNNSTTTPNNTQQSNTTPDQONQNPQAP"
misc_feature 1120..1122
              /gene="pbpla"
              /note="Active-site serine residue (Ser-370) that is
              acylated by penicillins and cephalosporins"
BASE COUNT   707 a 495 c 398 g 572 t
ORIGIN
Query Match      19.8%; Score 396; DB 1; Length 2172;
Best Local Similarity 99.1%; Pred. NO. 1.2e-211;
Matches 696; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 TAAATCTACGACATAAAATCAACTGCTGACTGGGTCTGAACGGCGGTCAA 60
    |||||||
DB 171 TAAATCTACGACATAAAATCAACTGCTGACTGGGTCTGAACGGCGGTCAA 230
    |||||||
QY 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
    |||||||
DB 231 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 290
    |||||||
QY 121 TCGCTTCTTCGACCACAGGGGATTGATACCATCCGTATCCTGGGAGCTTCTTCGCGCAA 180
    |||||||
DB 291 TCGCTTCTTCGACCACAGGGGATTGATACCATCCGTATCCTGGGAGCTTCTTCGCGCAA 350
    |||||||
QY 181 TCTGCAAGCAATTCCTCCAGGTGGATCACTCTCAACCCCAACAGTTGATTAAGTTGAC 240
    |||||||
```

```
Db 351 TCTGCAAGCAATTCCTCCAGGTGGATCAACTCTCAACCAACAGTTGATTAATGAC 410
QY 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
    |||||||
Db 411 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 470
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
    |||||||
Db 471 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 530
QY 361 GGTCTCATGTCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAACTACTATGGTAA 420
    |||||||
Db 531 GGTCTCATGTCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAACTACTATGGTAA 590
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGTGCTGGTGAATGCTCAGGCACC 480
    |||||||
Db 591 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGTGCTGGTGAATGCTCAGGCACC 650
QY 481 AAACCAATATGACCCCTATTTCACATCCAGAACGAGCCCAAGACCGCGGAACTTGTGCTT 540
    |||||||
Db 651 AAACCAATATGACCCCTATTTCACATCCAGAACGAGCCCAAGACCGCGGAACTTGTGCTT 710
QY 541 ATCTGAATGAAATAATCAAGGCTACATCTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
    |||||||
Db 711 ATCTGAATGAAATAATCAAGGCTACATCTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 770
QY 601 ACCAATCTGATGAGTACAAAGTCTCAAAATCAGCAAGTAATACCTCTTACATGGA 660
    |||||||
Db 771 ACCAATCTGATGAGTACAAAGTCTCAAAATCAGCAAGTAATACCTCTTACATGGA 830
QY 661 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGAAAC 702
    |||||||
Db 831 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGAAAC 872
RESULT 31
SPN403975
LOCUS      Streptococcus pneumoniae partial pbpla gene for penicillin binding
DEFINITION protein 1a, strain 960036.
ACCESSION  AJ403975
VERSION     AJ403975.1 GI:12055852
KEYWORDS    pbpla gene; penicillin binding protein 1a.
SOURCE      Streptococcus pneumoniae.
ORGANISM    Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
REFERENCE    1 (bases 1 to 489)
AUTHORS      Overweg,K., Bogaert,D., Sluijter,M., de Groot,R. and Hermans,P.W.M.
TITLE        Molecular characterization of Streptococcus pneumoniae
JOURNAL      penicillin-resistance in the Netherlands
AUTHORS      Unpublished
TITLE        Direct Submission
JOURNAL      Submitted (25-MAY-2000) Hermans P.W.M., Laboratory of Pediatrics,
              Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam,
              NETHERLANDS
FEATURES     Location/Qualifiers
  source      1..489
              /organism="Streptococcus pneumoniae"
              /strain="960036"
              /db_xref="taxon:1313"
  gene        1..489
              /gene="pbpla"
              /function="peptidoglycan transglycosylase-transpeptidase"
              /codon_start=1
              /transl_table=11
              /product="penicillin binding protein 1a"
              /protein_id="CAC20957.1"
              /db_xref="GI:12055853"
```


/translation="SNYPAYMDNLYKEVINOVETGYNLLTGTMDVYTNVDQEAQKH
LWDIYNTDEYVAYPDDELQVASTIVDSNGKVIAGLARGHOSNNVSFGINQAVETNRD
WGSTMKPITDYAPALEYGYVDSTATIVHDEPNYPCTNTPVYNWDRGYFGNITLQYAL
QOS"
BASE COUNT 154 a 117 c 98 g 120 t
ORIGIN

Query Match 19.4%; Score 387; DB 1; Length 489;
Best Local Similarity 99.6%; Pred. No. 1.4e-206;
Matches 487; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 638 AGTAATTACCCCTGCTTACATGATTAATACCTCAAGGAAGTCAATCAAGTTGAAGAA 697
|||||
Db 1 AGTAATTACCCCTGCTTACATGATTAATACCTCAAGGAAGTCAATCAAGTTGAAGAA 60
QY 698 GAAACAGGCTATAACCTACTACAACTGGGATGGATGCTTACACAAATGTAGACCAAGAA 757
|||||
Db 61 GAAACAGGCTATAACCTGCTCAACTGGGATGGATGCTTACACAAATGTAGACCAAGAA 120
QY 758 GCTCAAAAACATCTGTGGGATATTTACAATACAGACGAATAGCTTGCCCTATCCAGACGAT 817
|||||
Db 121 GCTCAAAAACATCTGTGGGATATTTACAATACAGACGAATAGCTTGCCCTATCCAGACGAT 180
QY 818 GAATTCGAAGTGCCTTCTACCATGTTGATGTTCTTAACGGTAAAGTCAATGCCCCAGCTA 877
|||||
Db 181 GAATTCGAAGTGCCTTCTACCATGTTGATGTTCTTAACGGTAAAGTCAATGCCCCAGCTA 240
QY 878 GGAGCAGCCATCAGTCAAGTAAATGTTCTTCGGGAATTAACCAAGCAGTAAACCAAAAC 937
|||||
Db 241 GGAGCAGCCATCAGTCAAGTAAATGTTCTTCGGGAATTAACCAAGCAGTAAACCAAAAC 300
QY 938 CGCGACTGGGGATCAACTATGAACCGATCACAGACTATGCTCCTGCTTGGAGTACGGT 997
|||||
Db 301 CGCGACTGGGGATCAACTATGAACCGATCACAGACTATGCTCCTGCTTGGAGTACGGT 360
QY 998 GTCTAGGATCAACTGCTACTATGTTTACAGTACGAGCCCTATAACTACCTGGGACAAAT 1057
|||||
Db 361 GTCTAGGATCAACTGCTACTATGTTTACAGTACGAGCCCTATAACTACCTGGGACAAAT 420
QY 1058 ACTCCTGTTTATAACTGGGATAGGGGCTACTTTGGCAACATCACCTTGAATACGCCCTG 1117
|||||
Db 421 ACCCTGTTTATAACTGGGATAGGGGCTACTTTGGCAACATCACCTTGAATACGCCCTG 480
QY 1118 CAACAAATCG 1126
|||||
Db 481 CAACAAATCG 489

RESULT 32
SPN403978 489 bp DNA linear BCT 06-JAN-2001
LOCUS Streptococcus pneumoniae partial.pbpla gene for penicillin binding
DEFINITION protein la, strain 950423.

ACCESSION AJ403978
VERSION AJ403978.1 GI:12055858
KEYWORDS pbpla gene; penicillin binding protein 1A.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae

REFERENCE 1 (bases 1 to 489)
AUTHORS Overweg,K., Bogaert,D., Sluifjter,M., de Groot,R. and Hermans,P.W.M.
TITLE Molecular characterization of Streptococcus pneumoniae

JOURNAL Unpublished
AUTHORS Hermans,P.W.M.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2000) Hermans P.W.M., Laboratory of Pediatrics,
Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam,
NETHERLANDS

FEATURES Location/Qualifiers

source 1..489
/organism="Streptococcus pneumoniae"
/strain="950423"
/db_xref="taxon:1313"
1..489
/gene="pbpla"
CDS
1..489
/gene="pbpla"
<1..489
/function="peptidoglycan transglycosylase-transpeptidase"
/codon_start=1
/transl_table=11
/product="penicillin binding protein 1a"
/protein_id="CAC20960.1"
/db_xref="GI:12055859"
/translation="SNYPAYMDNLYKEVINOVETGYNLLTGTMDVYTNVDQEAQKH
LWDIYNTDEYVAYPDDELQVASTIVDSNGKVIAGLARGHOSNNVSFGINQAVETNRD
WGSTMKPITDYAPALEYGYVDSTATIVHDEPNYPCTNTPVYNWDRGYFGNITLQYAL
QOS"
BASE COUNT 154 a 117 c 98 g 120 t
ORIGIN

Query Match 19.4%; Score 387; DB 1; Length 489;
Best Local Similarity 99.6%; Pred. No. 1.4e-206;
Matches 487; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 638 AGTAATTACCCCTGCTTACATGATTAATACCTCAAGGAAGTCAATCAAGTTGAAGAA 697
|||||
Db 1 AGTAATTACCCCTGCTTACATGATTAATACCTCAAGGAAGTCAATCAAGTTGAAGAA 60
QY 698 GAAACAGGCTATAACCTACTACAACTGGGATGGATGCTTACACAAATGTAGACCAAGAA 757
|||||
Db 61 GAAACAGGCTATAACCTGCTCAACTGGGATGGATGCTTACACAAATGTAGACCAAGAA 120
QY 758 GCTCAAAAACATCTGTGGGATATTTACAATACAGACGAATAGCTTGCCCTATCCAGACGAT 817
|||||
Db 121 GCTCAAAAACATCTGTGGGATATTTACAATACAGACGAATAGCTTGCCCTATCCAGACGAT 180
QY 818 GAATTCGAAGTGCCTTCTACCATGTTGATGTTCTTAACGGTAAAGTCAATGCCCCAGCTA 877
|||||
Db 181 GAATTCGAAGTGCCTTCTACCATGTTGATGTTCTTAACGGTAAAGTCAATGCCCCAGCTA 240
QY 878 GGAGCAGCCATCAGTCAAGTAAATGTTCTTCGGGAATTAACCAAGCAGTAAACCAAAAC 937
|||||
Db 241 GGAGCAGCCATCAGTCAAGTAAATGTTCTTCGGGAATTAACCAAGCAGTAAACCAAAAC 300
QY 938 CGCGACTGGGGATCAACTATGAACCGATCACAGACTATGCTCCTGCTTGGAGTACGGT 997
|||||
Db 301 CGCGACTGGGGATCAACTATGAACCGATCACAGACTATGCTCCTGCTTGGAGTACGGT 360
QY 998 GTCTAGGATCAACTGCTACTATGTTTACAGTACGAGCCCTATAACTACCTGGGACAAAT 1057
|||||
Db 361 GTCTAGGATCAACTGCTACTATGTTTACAGTACGAGCCCTATAACTACCTGGGACAAAT 420
QY 1058 ACTCCTGTTTATAACTGGGATAGGGGCTACTTTGGCAACATCACCTTGAATACGCCCTG 1117
|||||
Db 421 ACCCTGTTTATAACTGGGATAGGGGCTACTTTGGCAACATCACCTTGAATACGCCCTG 480
QY 1118 CAACAAATCG 1126
|||||
Db 481 CAACAAATCG 489

RESULT 33
SPN403978 2157 bp DNA linear BCT 11-MAY-1995
LOCUS S.pneumoniae (465) ponA gene for penicillin-binding protein 1a.
DEFINITION X67868
ACCESSION X67868
VERSION X67868.1 GI:47405
KEYWORDS penicillin-binding protein 1a; ponA gene.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

[illegible]

Streptococcus.

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 2157)
Hakenbeck, R.
Direct Submission
Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek.
Genetik, Ihnestr. 73, 1000 Berlin 33, FRG

2 (bases 1 to 2157)
Martin, C., Sibold, C. and Hakenbeck, R.
Relatedness of penicillin-binding protein la genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain
EMBO J. 11 (11), 3831-3836 (1992)

JOURNAL
MEDLINE
FEATURES

source
1. .2157
/organism="Streptococcus pneumoniae"
/strain="456"
/db_xref="taxon:1313"
/clone="pcMK1-pcMK16"
1. .2157
/gene="ponA"
1. .>2157
/gene="ponA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein la"
/protein_id="CAA48068.1"
/db_xref="GI:47406"
/translation="MNKPTILRLIKYLSISFLSLVIAIVLGGGVFFYYYSKAPSLSE
SKLVATTSKIYDNKNQIADIGSERVRNAQNDIPTDLVKAIYSIEDHRFFDHRGID
SIRILGALFNQLNSLOGGSTLTQOLIKLYFTSTSDOTISRKAQEWLAOLEOK
ATFKOELIAYINKVYMSNGNYGMQTAQONYGKDNLNLSLPALLAGMPAQNPQDP
YSHPEAAQRNRNLVLSEMKNQYISAEOYEKAVNTPIITDGLOSLKSASNPAYMDNVL
KEVINOVEQETGYNLLTGMVDYTNDOEAQHILWDIYNQVYSVPDDLEOVASTTV
DEVSNGKVAIQGLARHQASNVSGFNOAVETNRDGSAMKPITDYAPAIEGYVDSTAT
MYNDIPYNGPSTPYVWDRAYRFGNIITLQVALQOSNRVTAVETLNKVGLDRAKTFIN
GUGIDYPSMHYANALSSNTESNQYGASSERMAAAYAFAANGTTYKPMYIHKKVFS
DGSKEFNSVGTRAMKETATTMTDMKTILTYTGREGAYLPMLPOAGKTGTSNYTE
EVENHIKNITGYVAPDEMVEGYTRYKSYMAVWTGPVGDGFLVAALKVYSMTFY
LESDTHPEDWTPMDCGLFRNGBEVFKNRGARPITWRPSTQOSTAESSSDSSTSQS
STPTSTNNSTTNPNNTQQSNTPPDQNQPQAQ"

BASE COUNT 695 a 494 c 403 g 565 t

ORIGIN

Query Match 15.9%; Score 318; DB 1; Length 2157;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAAATCTACGACAATAAAAAATCAACTCATTCGTGACTTTGGGTTCTGAACGCCGCTCAA 60
|
|
|
Db 159 TAAAATCTAGCACAATAAAAAATCAACTCATTCGTGACTTTGGGTTCTGAACGCCGCTCAA 218
|
|
|
Qy 61 TGCCCAAGCTTAATGATATCCCACAGATTGGTTTAAGGCAATCGTTTCATCGAAGACCA 120
|
|
|
Db 219 TGCCCAAGCTTAATGATATCCCACAGATTGGTTTAAGGCAATCGTTTCATCGAAGACCA 278
|
|
|
Qy 121 TCAGTCTTCGACCACAGGGGATTCATACCATCCGATCCCTGGGAGCTTCTTCGCGAA 180
|
|
|
Db 279 TCGGTTCTTCGACCACAGGGGATTCATACCATCCGATCCCTGGGAGCTTCTTCGCGAA 338
|
|
|
Qy 181 TCTGCAAGAAGAAATTCCTCCAAGTGGATCAACTCTCACCCAACAGTGCATTAACTTGAC 240
|
|
|
Db 339 TCTGCAAGAAGTAATTCCTCCAAGTGGATCAACTCTCACCCAACAGTGCATTAACTTGAC 398
|
|
|
Qy 241 TTACTTTTTCAACTTCGACTTCCGACACAGACTATTTCTTCGTAAGGCTTCAGGAAGCTTGGTT 300
|
|
|
Db 399 TTACTTTTTCAACTTCGACTTCCGACACAGACTATTTCTTCGTAAGGCTTCAGGAAGCTTGGTT 458
|
|
|
Qy 301 AGCGATTTCAGTTAGACAAACAAACACCAAGCAAGAAATCTTTGACCTACTATATAATAA 360
|
|
|
Db 459 AGCGATTTCAGTTAGACAAACAAACCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA 518
|
|
|

GLGIDYPSMHYANAISSNTESNKQYGASSEKMAAAYAFANGIYHKPMYINKVYFVS
DSKKRFSVGTGRANKETAYMMTEMMKTVLAYGTGRGAYLPLWLAQAGKATQSNVTDD
ETEKHKTSYGVADPMDFVGYTRKYSMAVVTGYSNRLTPIVGDGFVAAKVYRSMITY
LSEDPHPEDWTPDGLFRNGEVEFKNGABSTWNSPAQOPPTSESSSSSDSTSQSS
STPSTNNSTTTDPNNNTQOSNTTDDQONQPPAPQ"

BASE COUNT 696 a 494 c 405 g 562 t
ORIGIN

Query Match 15.9%; Score 318; DB 1; Length 2157;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY	1	TAAATCTACGACAATAAAAAATCAACTCATCTGGTCTGGAACCGCGCTCAA	60
Db	159	TAAATCTACGACAATAAAAAATCAACTCATCTGGTCTGGAACCGCGCTCAA	218
QY	61	TGCCCAAGCTAATGATATTCCTCCACAGATTTGGTTAAGGCAATCGTTCTATCGAAGACCA	120
Db	219	TGCCCAAGCTAATGATATTCCTCCACAGATTTGGTTAAGGCAATCGTTCTATCGAAGACCA	278
QY	121	TCGCTTCTTCGACACAGGGGATGATACCATCCGTATCCTGGAGCTTCTTGCGCAA	180
Db	279	TCGCTTCTTCGACACAGGGGATGATACCATCCGTATCCTGGAGCTTCTTGCGCAA	338
QY	181	TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCAACAGTTGATTAAAGTTGAC	240
Db	339	TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCAACAGTTGATTAAAGTTGAC	398
QY	241	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	300
Db	399	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	458
QY	301	AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	360
Db	459	AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	518
QY	361	GGTCTACATCTTAATGGGAGATGGAATCGACAGACAGCTCAAACTACTATGGTAA	420
Db	519	GGTCTACATCTTAATGGGAGATGGAATCGACAGACAGCTCAAACTACTATGGTAA	578
QY	421	AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTCGCTGGGGAATGCTCAGGCACC	480
Db	579	AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTCGCTGGGGAATGCTCAGGCACC	638
QY	481	AAACCAATATGACCCCTATTTCACATCCAGAGCAGCCCAAGCCGCGAACTTGGTCTT	540
Db	639	AAACCAATATGACCCCTATTTCACATCCAGAGCAGCCCAAGCCGCGAACTTGGTCTT	698
QY	541	ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGATATGAGAAAGCAGTCAATAC	600
Db	699	ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGATATGAGAAAGCAGTCAATAC	758
QY	601	ACCAATTAAGTGGACTACAAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGGA	660
Db	759	ACCAATTAAGTGGACTACAAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGGA	818
QY	661	TAATTACTCTCAAGGA 675	
Db	819	TAATTACTCTCAAGGA 833	

RESULT 35
LOCUS SPPONA2 2157 bp DNA linear BCT 11-MAY-1995
DEFINITION S.pneumoniae (670) ponA gene for penicillin-binding protein 1a.
ACCESSION X67867
VERSION X67867.1 GI:47409
KEYWORDS penicillin-binding protein 1a; ponA gene.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae.
Bacteria: Firmicutes; Bacillus/Clostridium group: Streptococcaceae;
Streptococcus.

1 (bases 1 to 2157)
Hakenbeck, R.
Direct Submission
Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek.
Genetik, Imstr. 73, 1000 Berlin 33, FRG
2 (bases 1 to 2157)
Martin, C., Sibold, C. and Hakenbeck, R.
Relatedness of penicillin-binding protein 1a genes from different
clones of penicillin-resistant Streptococcus pneumoniae isolated in
South Africa and Spain
EMBO J. 11 (11), 3831-3836 (1992)
93010977 Location/Qualifiers
1..2157
/organism="Streptococcus pneumoniae"
/strain="670"
/db_xref="taxon:1313"
/clone="pcMK1-pcMK16"
1..2157
/gene="ponA"
1..>2157
/gene="ponA"
/codon_start=1
/transl_table=1
/product="penicillin-binding protein 1a"
/protein_id="CAA48067.1"
/db_xref="GI:47410"
/db_xref="SPTREMBL:Q54948"
/translation="MNKPIRLIKYLSIFSLVIAAIVLGGVFFYVSKAPLSLE
SKLVAATSSKIYDNKNQIADLGSERVAQANDIPTDLVKAIVSEHDFHRRGID
STRIGLAFRLNQSLSNGSTLTQILKLTFTSTSDQTTISRKAQEWLAQLQEQK
ATKOEITVYINKVYMSNGYMQTAQNYGKDLNLSLPQLALLAGMPAQNPQYDP
YSHPEAQQDRRLNLVLESEMNOGYISAEQYKAVNTPITDGLQSLKASNVPAYMDNVL
KVINQVEQETCYNLLTTCMDVYTNVDOEAQKHLMDIYNSDQXVSYVDDLOVASTV
DVSNGKVIQALGARHOASNVSGTNOAVETNRDWSGMRPITDYAIPAIEGYVDSTAT
MNDPIPNYPTGTSPIVNWDRAYFNITLQYALQQSRNVAVETLNKNGLDRAKTFLN
GLGIDYPSMHYANAISSNTESNKQYGASSEKMAAAYAFANGIYHKPMYINKVYFVS
DSKKRFSVGTGRANKETAYMMTEMMKTVLAYGTGRGAYLPLWLAQAGKATQSNVTDD
ETEKHKTSYGVADPMDFVGYTRKYSMAVVTGYSNRLTPIVGDGFVAAKVYRSMITY
LSEDPHPEDWTPDGLFRNGEVEFKNGABSTWNSPAQOPPTSESSSSSDSTSQSS
STPSTNNSTTTDPNNNTQOSNTTDDQONQPPAPQ"

BASE COUNT 696 a 494 c 405 g 562 t
ORIGIN

Query Match 15.9%; Score 318; DB 1; Length 2157;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY	1	TAAATCTACGACAATAAAAAATCAACTCATCTGGTCTGGAACCGCGCTCAA	60
Db	159	TAAATCTACGACAATAAAAAATCAACTCATCTGGTCTGGAACCGCGCTCAA	218
QY	61	TGCCCAAGCTAATGATATTCCTCCACAGATTTGGTTAAGGCAATCGTTCTATCGAAGACCA	120
Db	219	TGCCCAAGCTAATGATATTCCTCCACAGATTTGGTTAAGGCAATCGTTCTATCGAAGACCA	278
QY	121	TCGCTTCTTCGACACAGGGGATGATACCATCCGTATCCTGGAGCTTCTTGCGCAA	180
Db	279	TCGCTTCTTCGACACAGGGGATGATACCATCCGTATCCTGGAGCTTCTTGCGCAA	338
QY	181	TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCAACAGTTGATTAAAGTTGAC	240
Db	339	TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCAACAGTTGATTAAAGTTGAC	398
QY	241	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	300
Db	399	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	458
QY	301	AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	360
Db	459	AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	518

QY 361 GGTCTACATGCTAATGGGAACATATGAATGCAGACAGAGCTCAAACTACTATATGGTAA 420
|||||
Db 519 GGTCTACATGCTAATGGGAACATATGAATGCAGACAGAGCTCAAACTACTATATGGTAA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGTAATGCTCAGGCACC 480
Db 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGTAATGCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTACATCTCAGAAGCAGCCCAAGACCGCGGAAACTTGGTCTT 540
Db 639 AAACCAATATGACCCCTATTACATCTCAGAAGCAGCCCAAGACCGCGGAAACTTGGTCTT 698
QY 541 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 699 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 758
QY 601 ACCAATTACTGATGGACTACAAGTCTCAAACTCAGCAAGTAAATACCCCTGCTTACATGGA 660
Db 759 ACCAATTACTGATGGACTACAAGTCTCAAACTCAGCAAGTAAATACCCCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
Db 819 TAATTACCTCAAGGA 833

RESULT 36
SPON44
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

SPON44 2157 bp DNA linear BCT 11-MAY-1995
S.pneumoniae (56742) ponA gene for penicillin-binding protein 1a.
X67869
X67869.1 GI:47411
penicillin-binding protein 1a; ponA gene.
Streptococcus pneumoniae.
Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 2157)
Hakenbeck,R.
Direct Submission
Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek.
Genetik, Ihnestr. 73, 1000 Berlin 33, FRG
2 (bases 1 to 2157)
Martin,C., Sibold,C. and Hakenbeck,R.
Relatedness of penicillin-binding protein 1a genes from different
clones of penicillin-resistant Streptococcus pneumoniae isolated in
South Africa and Spain
EMBO J. 11 (11), 3831-3836 (1992)
93010977 Location/Qualifiers
1..2157
/organism="Streptococcus pneumoniae"
/strain="56742"
/db_xref="taxon:1313"
/clone="pcMK1-pcMK16"
1..2157
/gene="ponA"
1..>2157
/gene="ponA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="CAA48069.1"
/db_xref="GI:47412"
/db_xref="SPTREMBL:Q54949"
/translation="MNKPTILRLIKYLSFLVIAAIVLGGGVFFYVSKAPSLSE
SKLVATTSKLYDNKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGID
SIRILGAEIRNLQSNLSGGSTLTQOLIKLYFSTSDQTSRAQRAWLAIQLEQK
ATKQELIYIINKYVMSNGYGMQTAAQNYGKDLNLSLPQALLAGMPQAPNPDY
YSHPEAQDRNRNLVSEMKNGYISAEQYKAVNPTITDGLQSLKASNYPAYMDNYL
KEVINQVEQETGYNLLTGMVYTNVDOEAKHLWDIYNSDQYVSYPDLEQVASTV
DVSNGKVIQALGARHQASNVSGTNGQAVETNRDGSAMKPTIDYAPALEYGYDSTAT
MWNDIPYNPGTSTPYVNDRAYFGNITLQVALQOSRNVAVETLUNKVLQDRAKFTLN
GLGIDYPSMHYANAISNTTESNKYASSBKMAAAYAAAFANGGTYYPKMYIHKVYFS

BASE COUNT 696 a 494 c 403 g 564 t
ORIGIN
DGSKEFSNVGTRAMKETTAYMMTDMKTVLTYGTGRGAYLPWLPOAKTGTSNVTDE
EVENHKNTGVYAPDEMFGVTRKYSMVWVTGYSNRLTPIVGDGFLVAAKYVRSMITY
LSEDPEDWMPDGLFRNGEVEFKNGARPIWTEPSTQSSSTAESSSSDSSTSQSS
STTPSTNNSTTTPNPNNTQOSNTTTPDOQNPNPPAP"

Query Match 15.9%; Score 318; DB 1; Length 2157;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAAATCTAGGACATAAATACTCACTATGCTGACTTGGGTTCTGAAGCGCGGCTCAA 60
Db 159 TAAATCTAGGACATAAATACTCACTATGCTGACTTGGGTTCTGAAGCGCGGCTCAA 218
QY 61 TGCCCAAGCTAATGATATCCACACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 219 TGCCCAAGCTAATGATATCCACACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
QY 121 TCGCTTCTTCGACCCACAGGGGGATTGATACCATCCGTATCCTGGGAGCTTCTTTCGCGCA 180
Db 279 TCGCTTCTTCGACCCACAGGGGGATTGATTCATCCATCCGTATCCTGGGAGCTTCTTTCGCGCA 338
QY 181 TCTGCAAGCAATTCCTCCAAAGGTGGATCACTCTACCCCAACAGTTGATTAAGTTGAC 240
Db 339 TCTGCAAGCAATTCCTCCAAAGGTGGATCACTCTACCCCAACAGTTGATTAAGTTGAC 398
QY 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 399 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAAATAA 360
Db 459 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAAATAA 518
QY 361 GGTCTACATGCTAATGGGAACATATGGAATGCAGACAGAGCTCAAACTACTATATGGTAA 420
Db 519 GGTCTACATGCTAATGGGAACATATGGAATGCAGACAGAGCTCAAACTACTATATGGTAA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGTAATGCTCAGGCACC 480
Db 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGTAATGCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTACATCTCAGAAGCAGCCCAAGACCGCGGAAACTTGGTCTT 540
Db 639 AAACCAATATGACCCCTATTACATCTCAGAAGCAGCCCAAGACCGCGGAAACTTGGTCTT 698
QY 541 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 699 ATCTGAATGAATAATCAAGGTTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 758
QY 601 ACCAATTACTGATGGACTACAAGTCTCAAACTCAGCAAGTAAATACCCCTGCTTACATGGA 660
Db 759 ACCAATTACTGATGGACTACAAGTCTCAAACTCAGCAAGTAAATACCCCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
Db 819 TAATTACCTCAAGGA 833

RESULT 37
AX111407
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

AX111407 2157 bp DNA linear PAT 30-APR-2001
Sequence 2140 from Patent WO0123604.
AX111407
AX111407.1 GI:13927699
Streptococcus pneumoniae.
Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 2157)

AUTHORS Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M., Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers for detection of microorganisms
JOURNAL Patent: WO 0123604-A 2140 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
FEATURES Location/Qualifiers
source
1..2157
/organism="Streptococcus pneumoniae"
/strain="670"
/db_xref="taxon:1313"
BASE COUNT 696 a 494 c 405 g 562 t
ORIGIN
Query Match 15.9%; Score 318; DB 6; Length 2157;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 TAAATCTACGCAATAAATAACAATCACTATGCTGACTTGGTCTTGAACGCCGCGTCAA 60
Db 159 TAAATCTACGCAATAAATAACAATCACTATGCTGACTTGGTCTTGAACGCCGCGTCAA 218
QY 61 TGGCCCAAGCTAATGATATTCGCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 219 TGGCCCAAGCTAATGATATTCGCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
QY 121 TCGCTTCTTCGACCACAGGGGATTCATACCATCCGATCTCTCGGGAGCTTCTTCGGCAA 180
Db 279 TCGCTTCTTCGACCACAGGGGATTCATACCATCCGATCTCTCGGGAGCTTCTTCGGCAA 338
QY 181 TCTGCAAGCAATTCCTCCCAAGTGGATCACTCTCAACCCACAGTTGATTAAGTTGAC 240
Db 339 TCTGCAAGCAATTCCTCCCAAGTGGATCACTCTCAACCCACAGTTGATTAAGTTGAC 398
QY 241 TTACTTTTCAACTTCGACHTCCGACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 300
Db 399 TTACTTTTCAACTTCGACHTCCGACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 458
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGGACCTACTATATAATAA 360
Db 459 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGGACCTACTATATAATAA 518
QY 361 GGCTCTACATGCTTAATGGGAATATGGAATGCAGACAGAGCTCAAAACTACTATGGTAA 420
Db 519 GGCTCTACATGCTTAATGGGAATATGGAATGCAGACAGAGCTCAAAACTACTATGGTAA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGCTGGAATGCCTCAGGCACC 480
Db 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGCTGGAATGCCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTTCACATCCAGAAGCAGCCCAAGACCGCCGAAACTTGGTCTT 540
Db 639 AAACCAATATGACCCCTATTTCACATCCAGAAGCAGCCCAAGACCGCCGAAACTTGGTCTT 598
QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 699 ATCTGAAATGAAATCAAGGTTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 758
QY 601 ACCAATTACTGTGACTACAAGTCTCAATCAGCAAGTAATACCCCTGCTTACATGGA 660
Db 759 ACCAATTACTGTGACTACAAGTCTCAATCAGCAAGTAATACCCCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
Db 819 TAATTACCTCAAGGA 833
RESULT 38
AFI39883
LOCUS Streptococcus pneumoniae strain SP-665 penicillin-binding protein
DEFINITION la (pbpla) gene, complete cds.

AFI39883
AFI39883.1 GI:5410456
VERSION
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
TITLE Serotype 14 variants of the Spanish penicillin-resistant serotype 9V clone of Streptococcus pneumoniae arose by large recombinational replacements of the cpsA-pbpla region
JOURNAL Microbiology 145 (Pt 8), 2023-2031 (1999)
PUBMED 99392464
10463168
REFERENCE 2 (bases 1 to 2160)
AUTHORS Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-1999) The Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford, Oxfordshire OX1 3PS, UK
FEATURES Location/Qualifiers
source
1..2160
/organism="Streptococcus pneumoniae"
/strain="Sp-665"
/serotype="9V"
/db_xref="taxon:1313"
/note="Spanish member of the major penicillin-resistant Spanish serotype 9V clone"
gene
1..2160
/gene="pbpla"
CDS
1..2160
/gene="pbpla"
/function="peptidoglycan transglycosylase-transpeptidase"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="AAD43066.1"
/db_xref="GI:5410457"
/translation="MNKPTILRLKYLISFLISLVIAAIVLGGVFFVYVSKAPLSLSE
SKLVATTSKIIDNKNQLIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFDRGID
SIRLIGAFRLNLSNGSGSTLTQOLIKLTFTSTSDQTSIRKAQAWLAQLDEQK
ATQELITYINKVSNNGNYGMQTAQNYGYKDLNLSLPQLALLAGMAPNPQYDP
YSHPEAQDRRLNVLSEKNQGYISAEQYKAVNPTITDGLQSKASNYFAYMDNYL
KEVINOVEQETGYNLLTTGMDVYTNVDOEAQKHLMDIYVQYVSPDDDLQVASTV
DVSNCKVIAQLGARHQAQSNVSEFTNQAVETNRDWSANKPTIDYAPAYEYVDSTAT
MYNDIPYNYPGTSPFYVNDRAVFNITLQYALQOSRVNVAETLNKVLDRAKTFLN
GLGIDIPSMHYANATSSNTESKOYGASSERMAAAYAAFANGGTYIKPMYIHKVYFS
DGSKKFSNVGTRAMKETAYMTDMKTVLTYGTGRGAYLPWLPGQKGTSTNYTDE
EVENHIKNTGYVAPDEMFVGYTRKYSAVWTPYVSNRLPTIVGDGFLVAAKYRSMITY
LSEDTHPEDWTMPDGLFRNGEYFKNGARPIWTPESTQOSSSTAESSSSSSSTQSS
STTPSTNNSTTTNNNTQQSNTTDDQONQPQAP"
BASE COUNT 698 a 494 c 403 g 565 t
ORIGIN
Query Match 15.9%; Score 318; DB 1; Length 2160;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 TAAATCTACGCAATAAATAACAATCACTATGCTGACTTGGTCTTGAACGCCGCGTCAA 60
Db 159 TAAATCTACGCAATAAATAACAATCACTATGCTGACTTGGTCTTGAACGCCGCGTCAA 218
QY 61 TGGCCCAAGCTAATGATATTCGCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 219 TGGCCCAAGCTAATGATATTCGCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
QY 121 TCGCTTCTTCGACCACAGGGGATTCATACCATCCGATCTCTCGGGAGCTTCTTCGGCAA 180
Db 279 TCGCTTCTTCGACCACAGGGGATTCATACCATCCGATCTCTCGGGAGCTTCTTCGGCAA 338

QY	181	TCGCAAGCAATTCCTCCAGGTGGATCAACTCTCACCCAAACAGTTGATTAAGTTGAC	240
Db	339	TCGCAAGTAAATTCCTCCAGGTGGATCAACTCTCACCCAAACAGTTGATTAAGTTGAC	398
QY	241	TTACTTTTCAACTTCGACTTCCGACGACAGTATTTCTGTGAAGGCTCAGGAAGCTTGGTT	300
Db	399	TTACTTTTCAACTTCGACTTCCGACGACAGTATTTCTGTGAAGGCTCAGGAAGCTTGGTT	458
QY	301	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAAATAA	360
Db	459	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAAATAA	518
QY	361	GCTCTACATCTCTAATGGGAACATATGGAATGCAGACAGCAGCTCAAACTACTATGTTAA	420
Db	519	GCTCTACATCTCTAATGGCAACTATGGAATGCAGACAGCAGCTCAAACTACTATGTTAA	578
QY	421	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCCTGGCTGGTGAATGCCCTCAGGCACC	480
Db	579	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCCTGGCTGGTGAATGCCCTCAGGCACC	638
QY	481	AAACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGACCCGAAACTTGGTCTT	540
Db	639	AAACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGACCCGAAACTTGGTCTT	698
QY	541	ATCTGAATCAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC	600
Db	699	ATCTGAATCAAAAATCAAGGTTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC	758
QY	601	ACCAATTTACTGTAGGACTACAAAGTCTCAATCAGCAAGTAAATACCCTGCTTACATGGA	660
Db	759	ACCAATTTACTGTAGGACTACAAAGTCTCAATCAGCAAGTAAATACCCTGCTTACATGGA	818
QY	661	TAATTTACCTCAAGGA	675
Db	819	TAATTTACCTCAAGGA	833
RESULT 39			
AFI39884			
LOCUS	AFI39884	2160 bp	DNA linear BCT 26-AUG-1999
DEFINITION	Streptococcus pneumoniae strain PO-273 penicillin-binding protein		
ACCESSION	AFI39884	la (pbpla) gene, complete cds.	
VERSION	AFI39884.1	GI:5410458	
KEYWORDS	.		
SOURCE	Streptococcus pneumoniae.		
ORGANISM	Streptococcus pneumoniae		
REFERENCE	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.		
AUTHORS	1 (bases 1 to 2160)		
TITLE	Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.		
JOURNAL	Serotype 14 variants of the Spanish penicillin-resistant		
	9V clone of Streptococcus pneumoniae arose by large recombinational		
	replacements of the cpsA-pbpla region		
	Microbiology 145 (Pt 8), 2023-2031 (1999)		
MEDLINE	99392464		
PUBMED	10463168		
REFERENCE	2 (bases 1 to 2160)		
AUTHORS	Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-MAR-1999) The Wellcome Trust Centre for the		
	Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford, Oxfordshire OX1 3PS, UK		
FEATURES	Location/Qualifiers		
source	1..2160		
	/organism="Streptococcus pneumoniae"		
	/strain="PO-273"		
	/serotype="14"		
	/db_xref="taxon:1313"		
	/note="Polish variant of the major penicillin-resistant Spanish serotype 9V clone"		
gene	1..2160		

		/gene="pbpla"	
		1..2160	
		/gene="pbpla"	
		/function="peptidoglycan transglycosylase-transpeptidase"	
		/codon_start=1	
		/transl_table=11	
		/product="penicillin-binding protein la"	
		/protein_id="AAD43067.1"	
		/db_xref="GI:5410459"	
		/translation="MNKPTILRIKILISLSFLVIAAIVLGGGVFFYYVSKAPSLSE	
		SKLVATSSKIYDNKNOLIAIDGSERRVNAOANDIPTDLKAIYSIEDHFRFDHGRID	
		SIRLIGAFRLNLSQNSLOGGSTLQQLIKLYFSTSDOTISRKAQEAIAIOLEOK	
		ATKEILTYINKVMSNGYGMQTAQNYGKDLNLSLQALLAGMPQAPQNYDP	
		YSHPEAQDRLNLVSEMKNOGYSIAEQKAVNTPIIDGLOSLSKSNASNPAYMDNLY	
		KEVYNQVEQETGYNLLTGTMDVYVNDQEAQHLWDIYNSDYYSYPDDQLQVASTV	
		DVNSGKVIQALGRHQASVFGTNGAVETNRDWSAMKPTIDYAPATEYGVYSTAT	
		MVNDIPYNGTSTPYNMDRATFNTLQVALQOSRVNATVETLNLKVLGLDRAKTFLN	
		GLGIDYPSMHYANAISNTESNKQYGASSEKMAAAYAAFAFGTYKPKMYHKVVF	
		DGSKEFSNVGTRAMKETATYMTDMKTVLTGYGTGRGATLPWLQAGKGTSTNYTDE	
		EVENHIKNTGYVADPEMFVYTRKYSMAVWTGYSNRLTPIVGDGFLVAAKYVRSMTY	
		LSEGSNPEDWNIPEGLYRNGEFVKNGASTWSSPAQPPSTESSSSSDSTSQSS	
		STTPTSTNNSTTNPNNNTQQSNTPDQONQNPQAPQ"	
BASE COUNT		693 a	494 c 407 g 566 t
ORIGIN			
		Query Match 15.9%; Score 318; DB 1; Length 2160;	
		Best Local Similarity 99.0%; Pred. No. 1.3e-167;	
		Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
QY	1	TAATAATCTACGACATAAAAAATCAACTCATTCGTGACTTGGGTTCTGAACGCGCGTCAA	60
Db	159	TAATAATCTACGACATAAAAAATCAACTCATTCGTGACTTGGGTTCTGAACGCGCGTCAA	218
QY	61	TGCCCAAGCTAATGATATTCACACAGATTTCGTTAAGCAATCGTTCTATCGAAGACCA	120
Db	219	TGCCCAAGCTAATGATATTCACACAGATTTCGTTAAGCAATCGTTCTATCGAAGACCA	278
QY	121	TGCGTTCTTCGACCACAGGGGATTGATACCATCGTATCTCTGGAGCTTCTTCGCGCAA	180
Db	279	TGCGTTCTTCGACCACAGGGGATTGATTCATCCATCCGTATCTCTGGAGCTTCTTCGCGCAA	338
QY	181	TCTGCAAAAGCAATTCCTCCAAAGTGGATCAACTCTCACCCAAACAGTTGATTAAGTTGAC	240
Db	339	TCTGCAAAAGCAATTCCTCCAAAGTGGATCAACTCTCACCCAAACAGTTGATTAAGTTGAC	398
QY	241	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAGGCTCAGGAAGCTTGGTT	300
Db	399	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAGGCTCAGGAAGCTTGGTT	458
QY	301	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAAATAA	360
Db	459	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAAATAA	518
QY	361	GCTCTACATGCTCTAATGGAACTATGGAATGCAGACAGCTCAAACTACTATATGGTAA	420
Db	519	GCTCTACATGCTCTAATGGCAACTATGGAATGCAGACAGCTCAAACTACTATATGGTAA	578
QY	421	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCCTGGCTGGTGAATGCCCTCAGGCACC	480
Db	579	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCCTGGCTGGTGAATGCCCTCAGGCACC	638
QY	481	AAACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGACCCCGAAACTTGGTCTT	540
Db	639	AAACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGACCCCGAAACTTGGTCTT	698
QY	541	ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC	600
Db	699	ATCTGAAATGAAATCAAGGTTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC	758
QY	601	ACCAATTTACTGTAGCTACAAAGTCTCAAAATCAGCAAGTAAATACCCTGCTTACATGGA	660
Db	759	ACCAATTTACTGTAGCTACAAAGTCTCAAAATCAGCAAGTAAATACCCTGCTTACATGGA	818

University of Oxford, South Parks Road, Oxford, Oxfordshire OX1 3PS, UK

FEATURES
source

Location/Qualifiers
1. .2160
/organism="Streptococcus pneumoniae"
/strain="PO-342"
/serotype="14"
/db_xref="taxon:1313"
/note="Polish variant of the major penicillin-resistant Spanish serotype 9V clone"
1. .2160
/gene="pbpla"
1. .2160
/gene="pbpla"
/function="peptidoglycan transglycosylase-transpeptidase"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="A043069.1"
/db_xref="GI:5410463"

gene
CDS

BASE COUNT 693 a 494 c 407 g 566 t

ORIGIN

Query Match 15.9%; Score 318; DB 1; Length 2160;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAAATCTAGCAATAAATAAATCAATCATCTGCTGACTTGGGTCTGAAAGCCGCGTCAA 60
Db 159 TAAATCTAGCAATAAATAAATCAATCATCTGCTGACTTGGGTCTGAAAGCCGCGTCAA 218
QY 61 TGCCCAAGCTAATGATATCCACAGATTGGTTAAGCAATCGTTCTATCGAAGACCA 120
Db 219 TGCCCAAGCTAATGATATCCACAGATTGGTTAAGCAATCGTTCTATCGAAGACCA 278
QY 121 TCGCTTCTCGACACAGGGGGATTGATACCATCCGATCCGTTGGAGCTTCTTGGCGCAA 180
Db 279 TCGCTTCTCGACACAGGGGGATTGATACCATCCGATCCGTTGGAGCTTCTTGGCGCAA 338
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCACCACACAGTTGATTAAAGTTGAC 240
Db 339 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCACCACACAGTTGATTAAAGTTGAC 398
QY 241 TTACTTTTCAACTTCGACTTCCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 399 TTACTTTTCAACTTCGACTTCCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458
QY 301 AGCGATTTCAGTTAGACAAAGACCAAGCAAGAAATCTTGACCTTACCTATATAATAA 360
Db 459 AGCGATTTCAGTTAGACAAAGACCAAGCAAGAAATCTTGACCTTACCTATATAATAA 518
QY 361 GGTCTACATCTTAATGGGAATCTGGAATGCGACAGAGCTCAAACTACTATGGTAA 420
Db 519 GGTCTACATCTTAATGGGAATCTGGAATGCGACAGAGCTCAAACTACTATGGTAA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGAATGCTCAGGCACC 480
Db 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGAATGCTCAGGCACC 638

QY 481 AAACCAATATGACCCCTATTACATCCAGAGAGCCCAAGAGCCGCGAAACCTTGGTCTT 540
Db 639 AAACCAATATGACCCCTATTACATCCAGAGAGCCCAAGAGCCGCGAAACCTTGGTCTT 698
QY 541 ATCTGAATGAAATCAAGCTACATCTCTGCTGACACATATGAGAAACGACGTCATATC 600
Db 699 ATCTGAATGAAATCAAGCTACATCTCTGCTGACACATATGAGAAACGACGTCATATC 758
QY 601 ACCAAATCTGATGGACTACAAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGGA 660
Db 759 ACCAAATCTGATGGACTACAAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGGA 818
QY 661 TAATTACTCTCAAGGA 675
Db 819 TAATTACTCTCAAGGA 833

RESULT 42
AF139887
LOCUS
DEFINITION
2160 bp DNA linear BCT 26-AUG-1999
Streptococcus pneumoniae strain M134 penicillin-binding protein 1a (pbpla) gene, complete cds.
ACCESSION AF139887
VERSION AF139887.1 GI:5410464
KEYWORDS
SOURCE
ORGANISM
Streptococcus pneumoniae.
Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 2160)
Coffey, T.J., Daniels, M., Enright, M.C. and Spratt, B.G.
Serotype 14 variants of the Spanish penicillin-resistant serotype 9V clone of Streptococcus pneumoniae arose by large recombinational replacements of the cpsA-pbpla region
Microbiology 145 (Pt 8), 2023-2031 (1999)
99392464
10463168
REFERENCE 2 (bases 1 to 2160)
Coffey, T.J., Daniels, M., Enright, M.C. and Spratt, B.G.
AUTHORS
TITLE
Direct Submission
Submitted (31-MAR-1999) The Wellcome Trust Centre for the
Epidemiology of Infectious Disease, Department of Zoology,
University of Oxford, South Parks Road, Oxford, Oxfordshire OX1 3PS, UK

FEATURES
source
Location/Qualifiers
1. .2160
/organism="Streptococcus pneumoniae"
/strain="M134"
/serotype="14"
/db_xref="taxon:1313"
/note="Danish variant of the major penicillin-resistant Spanish serotype 9V clone"
1. .2160
/gene="pbpla"
1. .2160
/gene="pbpla"
/function="peptidoglycan transglycosylase-transpeptidase"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="A043070.1"
/db_xref="GI:5410465"

gene
CDS
translation="MNKPTILRLIKYLSIFLSLVIAAIVLGGVFYFVSKAPSLSE
SKLVATSSKIYDNKNQLIADLGSERRVNAQNDIPTDLVKAIVSIEDHRFDHIGID
SRIILGAFRLNQSLSGGSTLTQOLIKLITVFTSTDSQTSRKAQEAWLATLEQK
ATKQELITYYINKVYMSNGYGMQTAQNYGKDLNLSLPQALLAGNPQPNQYDP
YSHPEAAQRRLNLVSEMKNQGTISAEQYKAVNPITDGLQSLKASNPATMDNIL
KEVINQVETGYNLLTGMVYTNVDQAQHLNDIYNSDQYVSPDDQLQVASTV
DYSNGKVIAGLARHOASNVFCTNQAVETNRDWSAMKPTIDYAPAEIYVDSTAT
MYNDIPYNGTSTPVYNWDRVAFGNITLOYALQOSRNVTAETLNKVLGDRKTFLN
GLGIDYPSMHYANAISSNTTESKQYGASSEKMAAAYAFANGTGYKPMYTHKVYFS
DGSKEFSNVGTRANKETAYMMTDMKTVLIVYIGRGRAYLPWLPOAGTKGTSNYTDE
EVENHIKNTGYVAPDEMFGYTRKYSMAVNTGYSNRLTPIVGDFLVAAKVYRSMYTY
LSEGNPDNIEGKNGRFEVFNKARSTWSSPAQPPPTSESSSSSDSSSTSSQSS
STPTSTNNSTTTNPNNTTQOSNTTDPQONQNPQAP"


```

LSEDPHEDWTHPDGLFRNGEFVFNKARSTWNSPAPOQPPSTESSSSSDSSTSQSS
STPTSTNNSTTTNNNTQQSNTTDPQQNQNPQAP"
BASE COUNT      696 a  495 c  402 g  567 t
ORIGIN

Query Match      15.9%; Score 318; DB 1; Length 2160;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY  1  TAAATCTAGCAATAAATAAATCAACTCATGCTGACTTGGGTCTCAAGCGCGGTCAA 60
Db  159 TAAATCTAGCAATAAATAAATCAACTCATGCTGACTTGGGTCTCAAGCGCGGTCAA 218
QY  61  TGCCCAAGCTAATGATATATCCCAACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA 120
Db  219 TGCCCAAGCTAATGATATATCCCAACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA 278
QY  121 TCGCTTCTTCGACACAGGGGGATTGATACCATCCGATCTCTGGGAGCTTCTTTCGCGCAA 180
Db  279 TCGCTTCTTCGACACAGGGGGATTGATACCATCCGATCTCTGGGAGCTTCTTTCGCGCAA 338
QY  181 TCTGCAAGCAATTCCTCCCAAGGTGGATCACTCTCACCCACAGTTGATTAAGTTGAC 240
Db  339 TCTGCAAGCAATTCCTCCCAAGGTGGATCACTCTCACCCACAGTTGATTAAGTTGAC 398
QY  241 TTACTTTTCAACTTCGACTTCCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db  399 TTACTTTTCAACTTCGACTTCCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458
QY  301 AGCGATTGATTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
Db  459 AGCGATTGATTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 518
QY  361 GGTCTACATCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAAACTACTATGATAA 420
Db  519 GGTCTACATCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAAACTACTATGATAA 578
QY  421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGATGGATCGCTCAGGACAC 480
Db  579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGATGGATCGCTCAGGACAC 638
QY  481 AAACCAATATGACCCCTATTACATCCAGACAGCCCAAGACCCGCGGAACTTGGTCTT 540
Db  639 AAACCAATATGACCCCTATTACATCCAGACAGCCCAAGACCCGCGGAACTTGGTCTT 698
QY  541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC 600
Db  699 ATCTGAAATGAAAAATCAAGGTTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC 758
QY  601 ACCAATTACTGATGGACTACAAAGTCTCAAATCAGCAAGTAATTAACCTGCTTACATGGA 660
Db  759 ACCAATTACTGATGGACTACAAAGTCTCAAATCAGCAAGTAATTAACCTGCTTACATGGA 818
QY  661 TAATTACCTCAAGGA 675
Db  819 TAATTACCTCAAGGA 833

RESULT  43
AF139888
LOCUS
DEFINITION
Streptococcus pneumoniae strain URU-E135 penicillin-binding protein
1a (pbpla) gene, complete cds.
ACCESSION
AF139888
VERSION
AF139888.1
KEYWORDS
GI:5410466
SOURCE
Streptococcus pneumoniae.
ORGANISM
Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE
1 (bases 1 to 2160)
Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
```

```

Serotype 14 variants of the Spanish penicillin-resistant serotype
9v clone of Streptococcus pneumoniae arose by large recombinational
replacements of the cpsA-pbpla region
Microbiology 145 (Pt 8), 2023-2031 (1999)
9939246.
10463168
2 (bases 1 to 2160)
Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
Direct Submission
Submitted (31-MAR-1999) The Wellcome Trust Centre for the
Epidemiology of Infectious Disease, Department of Zoology,
University of Oxford, South Parks Road, Oxford, Oxfordshire OX1
3PS, UK
Location/Qualifiers
1. .2160
/organism="Streptococcus pneumoniae"
/strain="URU-E135"
/serotype="14"
/db_xref="taxon:1313"
/note="Uruguayan variant of the major penicillin-resistant
Spanish serotype 9v clone"
1. .2160
/genre="pbpla"
1. .2160
/genre="pbpla"
/function="peptidoglycan transglycosylase-transpeptidase"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="AAB43071.1"
/db_xref="GI:5410467"
/translation="MNKPTILRLIKYLSIFLSLVIAIIVGGGVFFYYVSKAPLSLSE
SKLVATTSKIYDNKNQIADLGSERRVNAQNDIPTDLVKAIVSTEDHRFDFHRGID
STRILGALRNQSLQSGSTLQOLIKLTVSTSDTISRKAQEWLAQLBOK
ATKQELIYVINKVYMSNGYGTAAQNYCYGKDLNLSLPOLALLAGMPQAPNOYDP
YSHPARAQRRLNVLSEMKNOGYSIAEQYEKAVNPTITDLSLSKASNPAYMDNYL
KEVINQVEQETGYNLLITGMDVYINQDEAKHLMDIYNDSQIVSTPDDDLQVASTV
DYSNGKVAQGLARHQAQNSVSEGTNOAVETNRDWSAMKPTIDYAPAEIYGYDSTAT
MYNDIPYVPGTSTPYNWDRAVFNITLQALQSRNVTAIVETLNKVLDRKFTLN
GLGIDYPSMHVANAISNTTESKQYSGASSEKMAAAYAFANGVYVYKPMYIHKVVF
DGSKEFSNVGTRAMKETAYNMTDMKTVLTYGTGARGAYLPWLQAGTKGTSNYVDE
EVENHIKNTGYVAPDEMFGVYTRKYSMAVWTCYSNRLPIVGDGLVAAKYVRSMITY
LSEDPHEDWTHPDGLFRNGEFVFNKARSTWNSPAPOQPPSTESSSSSDSSTSQSS
STPTSTNNSTTTNNNTQQSNTTDPQQNQNPQAP"
BASE COUNT      696 a  495 c  402 g  567 t
ORIGIN

Query Match      15.9%; Score 318; DB 1; Length 2160;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY  1  TAAATCTAGCAATAAATAAATCAACTCATGCTGACTTGGGTCTCAAGCGCGGTCAA 60
Db  159 TAAATCTAGCAATAAATAAATCAACTCATGCTGACTTGGGTCTCAAGCGCGGTCAA 218
QY  61  TGCCCAAGCTAATGATATATCCCAACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA 120
Db  219 TGCCCAAGCTAATGATATATCCCAACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA 278
QY  121 TCGCTTCTTCGACACAGGGGGATTGATACCATCCGATCTCTGGGAGCTTCTTTCGCGCAA 180
Db  279 TCGCTTCTTCGACACAGGGGGATTGATACCATCCGATCTCTGGGAGCTTCTTTCGCGCAA 338
QY  181 TCTGCAAGCAATTCCTCCCAAGGTGGATCACTCTCACCCACAGTTGATTAAGTTGAC 240
Db  339 TCTGCAAGCAATTCCTCCCAAGGTGGATCACTCTCACCCACAGTTGATTAAGTTGAC 398
QY  241 TTACTTTTCAACTTCGACTTCCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db  399 TTACTTTTCAACTTCGACTTCCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458
QY  301 AGCGATTGATTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
```



```
|||||
Db 459 AGCGATTTCAGTTAGAACAAAGCAACCAAGGAAATCTTCACCTACTATATAATAA 518
QY 361 GGTCCTACATGCTCTAAATGGGAACATGATGCAAGACAGAGCTCAAACTACTATGGTAA 420
Db 519 GGTCCTACATGCTCTAAATGGGAACATGATGCAAGACAGAGCTCAAACTACTATGGTAA 578
QY 421 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTGGCTGGGATGCAATGCTCAGGCACC 480
Db 579 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTGGCTGGGATGCAATGCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTCACATCCAGAAGACCCCAAGACCCGCGAAACTTGGTCTT 540
Db 639 AAACCAATATGACCCCTATTCACATCCAGAAGACCCCAAGACCCGCGAAACTTGGTCTT 698
QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 699 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 758
QY 601 ACCAATTACTGATGGACTACAAGTCTCAAAATCAGCAAGTAAATACCTGCTTACATGGA 660
Db 759 ACCAATTACTGATGGACTACAAGTCTCAAAATCAGCAAGTAAATACCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
Db 819 TAATTACCTCAAGGA 833

RESULT 44
AF139889
LOCUS
DEFINITION
Streptococcus pneumoniae strain URU-E157 penicillin-binding protein
1a (pbpla) gene, complete cds.
ACCESSION
AF139889
VERSION
AF139889.1 GI:5410468
KEYWORDS
Streptococcus pneumoniae.
SOURCE
Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 2160)
Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
Serotype 14 variants of the Spanish penicillin-resistant serotype
9V clone of Streptococcus pneumoniae arose by large recombinational
replacements of the cpsA-pbpla region
Microbiology 145 (Pt 8). 2023-2031 (1999)
99392464
MEDLINE
10463168
PUBMED
REFERENCE
2 (bases 1 to 2160)
Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
Direct Submission
Submitted (31-MAR-1999) The Wellcome Trust Centre for the
Epidemiology of Infectious Disease, Department of Zoology,
University of Oxford, South Parks Road, Oxford, Oxfordshire OX1
3PS, UK
FEATURES
source
1..2160
/organism="Streptococcus pneumoniae"
/strain="URU-E157"
/serotype="14"
/db_xref="taxon:1313"
/note="Uruguayan variant of the major penicillin-resistant
Spanish serotype 9V clone"
1..2160
/gene="pbpla"
1..2160
/gene="pbpla"
/function="peptidoglycan transglycosylase-transpeptidase"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="AAD43072.1"
/db_xref="GI:5410469"
```

```
/translation="MNKPTILRIKLSISFLSVIAAIVLGGGVFFYVSKAPSLSE
SKIVATTSSKIYDNKNQLIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFDFHRIID
SIRILGAPLURNQSNLOGSSTLTQOLIKLITYESTSTDOTISRKAQEWLAIOLEOK
ATKQETLTVYINKVYMSNGYMGOTAAQNYGKDLNLSIPOLALLAGLQPKQNOYDP
YSHPEAAQDRRLNVLSEKMNQGIISAEQYKAVNTPIITDGLQSLSKASNPAYMNYL
KEVINQEQETGYNLLTTGMDVTYNDQEKHLWDIYNSDQSVSPDDDLQVASTV
DYSNKGVIAOLGARHOASVSEFGTNOAVETNRDWSAMKPIITDIYPAIEYGVYDSTAT
MYNDIPYNYGSTPYVNDRAVFGNITLQYALQOSRNVATVETLNKVGIDRAKTPLN
GSGKFEFSNVGTAMKETAYMNTDMKTVLTYGTGGRGAVLPWLPQAGTGTGSKNTDE
EVENHINKTYVAPDEMFVGYTRKYSMAVWTGYSNRLTPIVGDGELVAAKYVRSMTY
LSEDTHPEDTWPDLGRNFEVFNKNGARSTWNSPAPQPPPTSESSSSSDSSTSSOSS
STTPSTNTTTNNPNNTQOSNTTDPQONPQAP"
BASE COUNT 696 a 495 c 402 g 567 t
ORIGIN
```

```
Query Match 15.9%; Score 318; DB 1; Length 2160;
Best Local Similarity 99.0%; Pred. No. 1.3e-17;
Matches 668; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 1 TAAATCTACGCAATAAATAAATCAACTCATTGCTGACTTGGTTCGAAACGCCGCGTCAA 60
Db 159 TAAATCTACGCAATAAATAAATCAACTCATTGCTGACTTGGTTCGAAACGCCGCGTCAA 218
QY 61 TGCCCAAGCTAATGATATTCACACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 219 TGCCCAAGCTAATGATATTCACACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
QY 121 TCCTCTTCGACACACAGGGGATTCATACCATCCGATCCTGGAGCTTCTTGGCGCAA 180
Db 279 TCCTCTTCGACACACAGGGGATTCATACCATCCGATCCTGGAGCTTCTTGGCGCAA 338
QY 181 TCTGCAAGCAATTCCTCCCAAGTGGATCACTCACCACACAGTTGATTAAAGTTGAC 240
Db 339 TCTGCAAGCAATTCCTCCCAAGTGGATCACTCACCACACAGTTGATTAAAGTTGAC 398
QY 241 TTACTTTCAACTTCGACTTCCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 399 TTACTTTCAACTTCGACTTCCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458
QY 301 AGCGATTTCAGTTAGAACAAAGCAACCAAGCAAGAAATCTTCACCTACTATATAATAA 360
Db 459 AGCGATTTCAGTTAGAACAAAGCAACCAAGCAAGAAATCTTCACCTACTATATAATAA 518
QY 361 GGTCCTACATGCTCTAAATGGGAACATGCAATGCAAGACAGAGCTCAAACTACTATGTTAA 420
Db 519 GGTCCTACATGCTCTAAATGGGAACATGCAATGCAAGACAGAGCTCAAACTACTATGTTAA 578
QY 421 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTGGCTGGGATGCAATGCTCAGGCACC 480
Db 579 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTGGCTGGGATGCAATGCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTCACATCCAGAAGACCCCAAGACCCGCGAAACTTGGTCTT 540
Db 639 AAACCAATATGACCCCTATTCACATCCAGAAGACCCCAAGACCCGCGAAACTTGGTCTT 698
QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 699 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 758
QY 601 ACCAATTACTGATGGACTACAAGTCTCAAAATCAGCAAGTAAATACCTGCTTACATGGA 660
Db 759 ACCAATTACTGATGGACTACAAGTCTCAAAATCAGCAAGTAAATACCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
Db 819 TAATTACCTCAAGGA 833
```

```
RESULT 45
AF139890
LOCUS
```

AF139890 2160 bp DNA linear BCT 26-AUG-1999


```
DEFINITION Streptococcus pneumoniae strain URU-E159 penicillin-binding protein
ACCESSION AF139890
VERSION AF139890.1 GI:5410470
KEYWORDS
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
TITLE Serotype 14 variants of the Spanish penicillin-resistant serotype
9V clone of Streptococcus pneumoniae arose by large recombinational
replacements of the cpsA-pbp1a region
JOURNAL Microbiology 145 (Pt 8), 2023-2031 (1999)
MEDLINE 99392464
PUBMED 10463168
REFERENCE 2 (bases 1 to 2160)
AUTHORS Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
TITLE Direct Submission
SUBMITTED (31-MAR-1999) The Wellcome Trust Centre for the
Epidemiology of Infectious Disease, Department of Zoology,
University of Oxford, South Parks Road, Oxford, Oxfordshire OX1
3PS, UK
FEATURES
    source
    Location/Qualifiers
        1..2160
            /organism="Streptococcus pneumoniae"
            /strain="URU-E159"
            /serotype="14"
            /db_xref="taxon:1313"
        note="Uruguayan variant of the major penicillin-resistant
        Spanish serotype 9V clone"
        1..2160
            /gene="pbp1a"
            1..2160
            /gene="pbp1a"
            /function="peptidoglycan transglycosylase-transpeptidase"
            /codon_start=1
            /transl_table=11
            /product="penicillin-binding protein 1a"
            /protein_id="AAD43073.1"
            /db_xref="GI:5410471"
            /translation="MKNKPTILRLIKYLSIFLSLVIAAIVLGGGVFFYVSKAPLSLSE
            SKLVATTSKIYDNKNLIAGSERRVNAQANDIPTDLVRAIVSIEDHRRFFDHRGID
            SIRILGAFRLNLSQSGSTLTQOLIKLVYFSTSDQTSIRKAOAWLAIOLEQK
            ATKOBILYYINKVYMSNGYGMOTAAONYGKDLNLSLPQALLACGAPQAPNOYDP
            YSHPEAODRRLNLSKMGNOYISAEQYERKAVPTITDGLQSLKSNYPAYMDNVL
            KEVINQVEQYGNLLTGMDEVYTNVDOERKHLMDIYNQSDQYVSYVSDDLQVASTV
            DVSNGKVIQAQGLARHQASNVSGTNAQVETNRDMSAMKPTIDYAPALEYGVYDSTAT
            MVNDIPYNGTSTPVYNDRAYFNITLQYALQOSRVNATVETLNKYLGLDRKFTFLN
            GIGIDYPSMHYANALISSNTTESKQYGASSEKMAAAYAAAFANGGTYYPKPMYIHKVVF
            DSKKFEFNVCTRAMKETAYMTDMKTVLTGYTGRGAYLPWLPOACKTGTSNYTD
            EVENHIKNTGYVADPEMVGYTRKYSMAVWGYSNRLTPIVGDGFLVAAKYRSMITY
            LSETHPEDWTPDGLFRNGEFVKNGARSTWNSPAPQPPSTESSSSSSSSSTSSQSS
            STTSTNSNTTNNNTNQSNNTTDDQONQNPQAPQ"
        BASE COUNT 696 a 495 c 402 g 567 t
        ORIGIN

Query Match 15.9%; Score 318; DB 1; Length 2160;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATAAAATCAACTCATCTGCTACTTGGGTCTGACCGCGGTCAA 60
|||||
DB 159 TAAATCTACGACATAAAATCAACTCATCTGCTACTTGGGTCTGACCGCGGTCAA 218
|||||
QY 61 TGCCCAAGTAATGATATTCACAGATTGGTTAAGGCAATCGTTCTATCATCAAGACCA 120
|||||
DB 219 TGCCCAAGTAATGATATTCACAGATTGGTTAAGGCAATCGTTCTATCATCAAGACCA 278
|||||
QY 121 TCGCTTCTTCGACCACAGGGGATGATACCATCCGTATCCCTGGGAGCTTCTTCGGCAA 180
|||||
```

```
DB 279 TCGCTTCTTCGACCACAGGGGATGATCCATCCGTATCCCTGGGAGCTTCTTCGGCAA 338
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGGATCACTCTCAGCCCAACAGCTTGATTAAAGTTGAC 240
|||||
DB 339 TCTGCAAGTAATTCCTCCCAAGGTGGATCACTCTCAGCCCAACAGCTTGATTAAAGTTGAC 398
QY 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
|||||
DB 399 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
|||||
DB 459 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 518
QY 361 GGTCTACATGCTCTAATGGGAACCTATGGAATGCAGACAGCAGCTCAAAACTACTATATGATA 420
|||||
DB 519 GGTCTACATGCTCTAATGGGAACCTATGGAATGCAGACAGCAGCTCAAAACTACTATATGATA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGAATGCCTCAGGCACC 480
|||||
DB 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGAATGCCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTCACATCCAGACAGCCCAAGACCCCGGAACTTGTCTTT 540
|||||
DB 639 AAACCAATATGACCCCTATTCACATCCAGACAGCCCAAGACCCCGGAACTTGTCTTT 698
QY 541 ATCTGAATGAAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
|||||
DB 699 ATCTGAATGAAATAATCAAGGTTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 758
QY 601 ACCAATATGATGGACTACAAAGTCTCAAACTCAGCAAGTAATTAACCTGCTTACATGGA 660
|||||
DB 759 ACCAATATGATGGACTACAAAGTCTCAAACTCAGCAAGTAATTAACCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
|||||
DB 819 TAATTACCTCAAGGA 833

RESULT 46
AF159448
LOCUS Streptococcus pneumoniae penicillin-binding protein 1a (pbp1a)
DEFINITION gene, complete cds.
ACCESSION AF159448
VERSION AF159448.1 GI:5726354
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
REFERENCE 1 (bases 1 to 2160)
AUTHORS Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
TITLE Serotype 14 variants of the Spanish penicillin-resistant serotype
9V clone of Streptococcus pneumoniae arose by large recombinational
replacements of the cpsA-pbp1a region
JOURNAL Microbiology 145 (1999) In press
REFERENCE 2 (bases 1 to 2160)
AUTHORS Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
TITLE Direct Submission
SUBMITTED (15-JUN-1999) The Wellcome Trust Centre for the
Epidemiology of Infectious Disease, Department of Zoology,
University of Oxford, South Parks Road, Oxford, Oxfordshire OX1
3FY, UK
FEATURES
    source
    Location/Qualifiers
        1..2160
            /organism="Streptococcus pneumoniae"
            /strain="URU-E206"
            /serotype="14"
            /db_xref="taxon:1313"
            /country="Uruguay"
            /note="variant of the major penicillin-resistant Spanish
            serotype 9V clone"
```


Db	219	 TGCCAAAGCTAATGATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA	278
Qy	121	TCGCTTCTTCGACCACAGGGGATGATPACCATCCGTATCTCTGGGAGCTTCTTTCGCGAA	180
Db	279	TCGCTTCTTCGACCACAGGGGATGATPACCATCCGTATCTCTGGGAGCTTCTTTCGCGAA	338
Qy	181	TCTGAAAGCAATTCCTCCCAAGTGGGATCAACTCTCACCCACAGTTGATTAAGTTGAC	240
Db	339	TCTGAAAGTAAATCCCTCCCAAGTGGGATCAACTCTCACCCACAGTTGATTAAGTTGAC	398
Qy	241	TTACTTTTCAACTTCGACCTCCGACCACAGCTATTTCTCGTAGGCTCAGGAAGCTTGGTT	300
Db	399	TTACTTTTCAACTTCGACCTCCGACCACAGCTATTTCTCGTAGGCTCAGGAAGCTTGGTT	458
Qy	301	AGCGATTTCAGTTAGAACAAAAAGCAACACAGAAATCTTGACCTACTATATAAATAA	360
Db	459	AGCGATTTCAGTTAGAACAAAAAGCAACACAGAAATCTTGACCTACTATATAAATAA	518
Qy	361	GGTCTACATGCTAATGGGAACTATGGAATCGACAGCAGCTCAAAACTACTATGGTAA	420
Db	519	GGTCTACATGCTAATGGGAACTATGGAATCGACAGCAGCTCAAAACTACTATGGTAA	578
Qy	421	AGACCTCAATTAATTAAGTTTACCTTCAGTTAGCCTTGCTGGCTGGAATGCCTCAGGCAC	480
Db	579	AGACCTCAATTAATTAAGTTTACCTTCAGTTAGCCTTGCTGGCTGGAATGCCTCAGGCAC	638
Qy	481	AAACCAATATGACCCCTATTACATCCAGAGCAGCCCCAAGCCGCGAACTTGGTCTT	540
Db	639	AAACCAATATGACCCCTATTACATCCAGAGCAGCCCCAAGCCGCGAACTTGGTCTT	698
Qy	541	ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC	600
Db	699	ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC	758
Qy	601	ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAAGTAATTAACCTGCTTCATGGA	660
Db	759	ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAAGTAATTAACCTGCTTCATGGA	818
Qy	661	TAATTACCTCAAGGA	675
Db	819	TAATTACCTCAAGGA	833
RESULT	48		
LOCUS	AF210750	2160 bp DNA linear BCT 17-SEP-2001	
DEFINITION	Streptococcus pneumoniae isolate SP 1470 penicillin-binding protein 1A (pbp1a) gene, complete cds.		
ACCESSION	AF210750		
VERSION	AF210750.1	GI:6563346	
KEYWORDS			
SOURCE	Streptococcus pneumoniae.		
ORGANISM	Streptococcus pneumoniae		
REFERENCE	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.		
AUTHORS	1 (bases 1 to 2160)		
TITLE	Ferroni,A. and Berche,P. Alterations to penicillin-binding proteins 1A, 2B and 2X amongst penicillin-resistant clinical isolates of Streptococcus pneumoniae serotype 23F from the nasopharyngeal flora of children		
JOURNAL	J. Med. Microbiol. 50 (9), 828-832 (2001)		
MEDLINE	21432820		
PUBMED	11549185		
REFERENCE	2 (bases 1 to 2160)		
AUTHORS	Ferroni,A. and Berche,P.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-DEC-1999) Microbiology, Hospital		
FEATURES	Necker-Enfants-Malades, 149 rue de Sevres, Paris 75015, France		
source	Location/Qualifiers		
	1. .2160		
	/organism="Streptococcus pneumoniae"		
	/serotype="23F"		


```
|||||
Db 759 ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAAGTAATACCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
Db 819 TAATTACCTCAAGGA 833

RESULT 49
AF210751 AF210751 2160 bp DNA linear BCT 17-SEP-2001
LOCUS Streptococcus pneumoniae isolate SP 22861 penicillin-binding
DEFINITION protein 1A (pbpla) gene, complete cds.
ACCESSION AF210751
VERSION AF210751.1 GI:6563348
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 2160)
Ferroni,A. and Berche,P.
Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
penicillin-resistant clinical isolates of Streptococcus pneumoniae
serotype 23F from the nasopharyngeal flora of children
J. Med. Microbiol. 50 (9), 828-832 (2001)
21432820
PUBMED 11549185
2 (bases 1 to 2160)
Ferroni,A. and Berche,P.
Direct Submission
Submitted (03-DEC-1999) Microbiology, Hospital
Necker-Enfants-Malades, 149 rue de Sevres, Paris 75015, France
FEATURES
Source
1..2160
/organism="Streptococcus pneumoniae"
/serotype="23F"
/isolate="SP 22861"
/db_xref="taxon:1313"
1..2160
/gene="pbpla"
1..2160
/gene="pbpla"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1A"
/protein_id="AA17261.1"
/db_xref="GI:6563349"
/translation="MNKPTILRLIKYLSIFSLVIAIVLGGVFFYYVSKAPLSLSE
SKLVATTSKIYDNKNLIADLGSERRVNAQNDIPTDLVKAIVSDIEDHRFDHRIID
STRILGAFRLNLSQSGSLTQOLIKLTVFSTSDQTSRKAQEAWLAIQLEQK
ATKQELITYINKVYMSNGYCMQTAQNYGKDLNLSLPOLALLAGMPQAPNDYD
YSHPEAQRRLNLVSEKMGQYISAEQYKAVNPTITDGLSLKASNYPAYMDNLY
KEVINQFQETGYNLTGMDVITNVDAQRLMDIYNDSQIVSTPDDDLQVASTV
DVNSKVLQAARHOASNVSEFTNQAVETNRDWSAMKPTIDYAPAEIYGVYDSTAT
MYNDIPYAPGSTVYWNDRAYFGNITLQYALQSRNVYAVETLNKVLDRAKTFLN
GIGIDYPSHVNANVLSSTNSKQYKASKEKAAAYAFANGGIYHKPMYINKVYFS
DGSKFSDVGTFRAMKETTAYMMETMMKTVLAVGTCRGAYLPWLAQAGKTGTSNYTDD
ELEKHKITGYVAPDEMFGVYTRKYSMAVWTCYSNRLTPIVGDGLVAAKGVYRSMITY
LSEGNPNEDNTPGLIRNGEYFRNGARSTWSSPAPQPPSTSESSSSSSSSSTSQSN
STTPSTNNSTTTNNNTNQOQNTTDPQONQNPQAP"
BASE COUNT 700 a 491 c 405 g 564 t
ORIGIN

Query Match 15.9%; Score 318; DB 1; Length 2160;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAAATCTAGCACAATAAATAACACTCATGCTGACTGGGTCTGAAGCCGCGTCAA 60
|||||
Db 159 TAAATCTAGCACAATAAATAACACTCATGCTGACTGGGTCTGAAGCCGCGTCAA 218
```

```
QY 61 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGGAATCGTTTCTATCGAAGACCA 120
|||||
Db 219 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGGAATCGTTTCTATCGAAGACCA 278
QY 121 TCGCTTCTTCGACCACAGAGGGGATTTGATACCATCCGATCCGATCCTGGGAGCTTCTTGGCAA 180
|||||
Db 279 TCGCTTCTTCGACCACAGAGGGGATTTGATCCATCCGATCCTGGGAGCTTCTTGGCAA 338
QY 181 TCTGCAAGCAATTCCTCCCAAGTGATCAACTCTCACCAACAGTGTGATTAAAGTTGAC 240
|||||
Db 339 TCTGCAAGCAATTCCTCCCAAGTGATCAACTCTCACCAACAGTGTGATTAAAGTTGAC 398
QY 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
|||||
Db 399 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
|||||
Db 459 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 518
QY 361 GGTCTACATGCTTAATGGGAACCTATGGAATGCAGACAGAGCTCAAAACTACTATGTAA 420
|||||
Db 519 GGTCTACATGCTTAATGGGAACCTATGGAATGCAGACAGAGCTCAAAACTACTATGTAA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGCTGGAATGCTCAGGCACC 480
|||||
Db 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGCTGGAATGCTCAGGCACC 638
QY 481 AAACCATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACCGCGGAACCTTGGCTTT 540
|||||
Db 639 AAACCATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACCGCGGAACCTTGGCTTT 698
QY 541 ATCTGAAATGAAAAAATAAGCTACATCTCTGCTGAAACAGTATGAGAAAACAGTCAATAC 600
|||||
Db 699 ATCTGAAATGAAAAAATAAGCTACATCTCTGCTGAAACAGTATGAGAAAACAGTCAATAC 758
QY 601 ACCAATTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGGA 660
|||||
Db 759 ACCAATTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
Db 819 TAATTACCTCAAGGA 833
```

```
RESULT 50
AX110436 AX110436 2160 bp DNA linear PAT 30-APR-2001
LOCUS Sequence 1169 from Patent WO0123604.
DEFINITION AX110436
ACCESSION AX110436
VERSION AX110436.1 GI:13926728
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Bergeron,M.G., Bolissnot,M., Huletsky,A., m Nard,C., Ouellette,M.,
Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 1169 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
FEATURES
source
1..2160
/organism="Streptococcus pneumoniae"
/strain="SP-665"
/db_xref="taxon:1313"
BASE COUNT 698 a 494 c 403 g 565 t
ORIGIN

Query Match 15.9%; Score 318; DB 6; Length 2160;
```


Best Local Similarity 99.0%; Pred. No. 1.3e-167;			
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;			
QY	1	TAAATCTACGACAATAAATAAATCAACTCATTTGCTGACTTGGGTTCTGAACGCGCGTCAA	60
Db	159	TAAATCTACGACAATAAATAAATCAACTCATTTGCTGACTTGGGTTCTGAACGCGCGTCAA	218
QY	61	TGCCCCAAGCTAATGATATTCCTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA	120
Db	219	TGCCCCAAGCTAATGATATTCCTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA	278
QY	121	TCGCTTCTTCGACACAGGGGGATTGATACCATCCGTATCCCTGGGAGCTTCTTTCGGCCAA	180
Db	279	TCGCTTCTTCGACACAGGGGGATTGATTCATCCGTATCCCTGGGAGCTTCTTTCGGCCAA	338
QY	181	TCGCAAGCAATTCCTCCACGGTGGATCAACTCTCACCCACAGTTGATTAAGTTGAC	240
Db	339	TCGCAAGCAATTCCTCCACGGTGGATCAACTCTCACCCACAGTTGATTAAGTTGAC	398
QY	241	TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTTGGTT	300
Db	399	TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTTGGTT	458
QY	301	AGCGATTCACTTAGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	360
Db	459	AGCGATTCACTTAGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	518
QY	361	GGTCTACATCTTAATGGGAACATATGAATGCAGACAGAGCTCAAAACTTACTATGGTAA	420
Db	519	GGTCTACATCTTAATGGGAACATATGAATGCAGACAGAGCTCAAAACTTACTATGGTAA	578
QY	421	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGGCTGGGATGCCTCAGGCACC	480
Db	579	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGGCTGGGATGCCTCAGGCACC	638
QY	481	AAACCAATATGACCCCTATTTCACATCCAGACAGAGCCCGGAACTTGGCTTT	540
Db	639	AAACCAATATGACCCCTATTTCACATCCAGACAGAGCCCGGAACTTGGCTTT	698
QY	541	ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	600
Db	699	ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	758
QY	601	ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAAGTAATTACCTGCTTACATGGA	660
Db	759	ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAAGTAATTACCTGCTTACATGGA	818
QY	661	TAATTACCTCAAGGA	675
Db	819	TAATTACCTCAAGGA	833

Search completed: June 13, 2002, 12:36:59
Job time: 14246 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 10:52:32 ; Search time 325.44 Seconds
(without alignments)
10546.067 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATATGACATAAAA.....ATCCTCAACGACACCA 1999

Scoring table: OLIGO_NUC
Gap 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 15

Total number of hits satisfying chosen parameters: 3497

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N_Geneseq_032802.*

- 1: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	1999	19 AAV27323	Streptococcus pneu
2	1999	100.0	10711	19 AAV52278	Streptococcus pneu
3	1438	71.9	2048	17 AH08027	S. pneumoniae peni
4	1438	71.9	2160	22 AAH02055	Streptococcus pneu
5	1438	71.9	2160	23 AAS55688	Streptococcus pneu
6	1198	59.9	1249	22 AAH02067	Streptococcus pneu
7	1014	50.7	1260	20 AAH02067	Streptococcus pneu
8	940	47.0	1260	20 AAH02067	Streptococcus pneu
9	903	45.2	1260	20 AAH02067	Streptococcus pneu

10	903	45.2	1260	20 AAH02067	Streptococcus pneu
11	838	41.9	1260	20 AAH02067	Streptococcus pneu
12	824	41.2	1260	20 AAH02067	Streptococcus pneu
13	796	39.8	960	16 AAQ83241	Streptococcus pneu
14	764	38.2	1260	20 AAH02067	Streptococcus pneu
15	745	37.3	1260	20 AAH02067	Streptococcus pneu
16	620	31.0	930	22 AAH02067	Streptococcus pneu
17	590	29.5	1199	22 AAH02067	Streptococcus pneu
18	581	29.1	1260	20 AAH02067	Streptococcus pneu
19	567	28.4	1229	22 AAH02067	Streptococcus pneu
20	554	27.7	1260	20 AAH02067	Streptococcus pneu
21	541	27.1	2172	19 AAH02067	S. pneumoniae deri
22	396	19.8	2160	22 AAH02067	Streptococcus pneu
23	318	15.9	2157	22 AAH02067	Streptococcus pneu
24	318	15.9	2160	22 AAH02067	Streptococcus pneu
25	318	15.9	2160	22 AAH02067	Streptococcus pneu
26	266	13.3	1260	20 AAH02067	Streptococcus pneu
27	228	11.4	1195	22 AAH02067	Streptococcus pneu
28	209	10.5	1260	20 AAH02067	Streptococcus pneu
29	178	8.9	1212	22 AAH02067	Streptococcus pneu
30	178	8.9	1222	22 AAH02067	Streptococcus pneu
31	175	8.8	1220	22 AAH02067	Streptococcus pneu
32	159	8.0	1216	22 AAH02067	Streptococcus pneu
33	129	6.5	1201	22 AAH02067	Streptococcus pneu
34	127	6.4	1212	22 AAH02067	Streptococcus pneu
35	127	6.4	1223	22 AAH02067	Streptococcus pneu
36	126	6.3	1211	22 AAH02067	Streptococcus pneu
37	126	6.3	1218	22 AAH02067	Streptococcus pneu
38	126	6.3	1213	22 AAH02067	Streptococcus pneu
39	122	6.1	1214	22 AAH02067	Streptococcus pneu
40	121	6.1	1207	22 AAH02067	Streptococcus pneu
41	120	6.0	1212	22 AAH02067	Streptococcus pneu
42	115	5.8	1242	22 AAH02067	Streptococcus pneu
43	110	5.5	1223	22 AAH02067	Streptococcus pneu
44	104	5.2	1225	22 AAH02067	Streptococcus pneu
45	101	5.1	1260	20 AAH02067	Streptococcus pneu
46	101	5.1	1260	20 AAH02067	Streptococcus pneu
47	97	4.9	930	22 AAH02067	Streptococcus pneu
48	97	4.9	1260	20 AAH02067	Streptococcus pneu
49	90	4.5	782	22 AAH02067	Streptococcus pneu
50	90	4.5	813	22 AAH02067	Streptococcus pneu
51	73	3.7	1225	22 AAH02067	Streptococcus pneu
52	44	2.2	930	22 AAH02067	Streptococcus pneu
53	25	1.3	25	22 AAH02067	Streptococcus pneu
54	25	1.3	8395	20 AAH02067	Streptococcus pneu
55	24	1.2	24	22 AAH02067	Streptococcus pneu
56	24	1.2	24	22 AAH02067	Streptococcus pneu
57	23	1.2	33	19 AAH02067	Streptococcus pneu
58	22	1.1	22	22 AAH02067	Streptococcus pneu
59	22	1.1	22	22 AAH02067	Streptococcus pneu
60	22	1.1	2337	23 AAH02067	Streptococcus pneu
61	20	1.0	20	22 AAH02067	Streptococcus pneu
62	20	1.0	20	22 AAH02067	Streptococcus pneu
63	20	1.0	25	22 AAH02067	Streptococcus pneu
64	20	1.0	550	22 AAH02067	Streptococcus pneu
65	20	1.0	554	22 AAH02067	Streptococcus pneu
66	20	1.0	680	20 AAH02067	Streptococcus pneu
67	20	1.0	5124	21 AAH02067	Streptococcus pneu
68	20	1.0	5410	22 AAH02067	Streptococcus pneu
69	20	1.0	5766	22 AAH02067	Streptococcus pneu
70	20	1.0	7268	22 AAH02067	Streptococcus pneu
71	19	1.0	19	19 AAH02067	Streptococcus pneu
72	19	1.0	19	19 AAH02067	Streptococcus pneu
73	19	1.0	428	21 AAH02067	Streptococcus pneu
74	19	1.0	822	23 AAH02067	Streptococcus pneu
75	19	1.0	1295	21 AAH02067	Streptococcus pneu
76	19	1.0	2193	19 AAH02067	Streptococcus pneu
77	19	1.0	5448	19 AAH02067	Streptococcus pneu
78	19	1.0	23041	22 AAH02067	Streptococcus pneu
79	19	1.0	23135	22 AAH02067	Streptococcus pneu
80	19	1.0	23135	22 AAH02067	Streptococcus pneu
81	19	1.0	1830121	17 AAH02067	Streptococcus pneu
82	19	1.0	1830121	17 AAH02067	Streptococcus pneu

83	18	0.9	18	19	AAV64670	S. pneumoniae PBP2	156	18	0.9	240825	22	AAF24497	Human PG-3 gene.
84	18	0.9	18	19	AAV64671	S. pneumoniae PBP2	c 157	17	0.9	27	19	AAV27438	Streptococcus pneu
c 85	18	0.9	21	22	AAH01135	S. pneumoniae ppp1	158	17	0.9	27	20	AAV25035	Cyclin-dependent k
86	18	0.9	24	22	AAH01204	S. pneumoniae ppp1	159	17	0.9	27	20	AAV25029	Cyclin-dependent k
c 87	18	0.9	383	22	AAI91836	Human polynucleoti	c 160	17	0.9	35	21	AAZ98275	Mouse IP3 sponge g
88	18	0.9	439	22	AAI63993	Human polynucleoti	c 161	17	0.9	36	21	AAZ98274	Mouse IP3 sponge g
89	18	0.9	439	22	AAI63993	Human polynucleoti	c 162	17	0.9	99	21	AAI61615	Human secreted pro
90	18	0.9	461	23	AAI18695	Genomic sequence #	163	17	0.9	130	18	AAV77571	Staphylococcus aur
c 91	18	0.9	491	23	AAI18695	Drosophila melanog	164	17	0.9	130	18	AAV77571	Staphylococcus aur
92	18	0.9	512	23	AAI18695	DNA encoding novel	c 165	17	0.9	226	21	AAV41385	Human secreted exp
93	18	0.9	526	21	AAI18695	cDNA #1069 encodin	166	17	0.9	234	21	AAV41385	Human secreted exp
94	18	0.9	541	22	AAI18695	Arabidopsis thalia	c 167	17	0.9	306	21	AAV41385	Human secreted exp
95	18	0.9	541	22	AAI18695	Human foetal liver	c 168	17	0.9	332	22	AAI187203	Neisseria meningit
96	18	0.9	541	22	AAI18695	Probe #8173 for ge	c 169	17	0.9	340	21	AAI187203	Human polynucleoti
97	18	0.9	541	22	AAI18695	Human brain expres	c 170	17	0.9	341	21	AAI187203	Human secreted pro
98	18	0.9	541	22	AAI18695	Human bone marrow	c 171	17	0.9	341	21	AAI187203	Neisseria meningit
99	18	0.9	541	22	AAI18695	Probe #7374 for ge	c 172	17	0.9	402	21	AAI187203	Neisseria meningit
c 100	18	0.9	541	22	AAI18695	Probe #11033 used	c 173	17	0.9	417	22	AAI187203	Human polynucleoti
c 101	18	0.9	612	20	AAI18695	Prostate cancer as	c 174	17	0.9	462	22	AAI187203	Human polynucleoti
c 102	18	0.9	700	21	AAI18695	Human inflammatory	c 175	17	0.9	500	23	AAI187203	DNA encoding novel
c 103	18	0.9	724	21	AAI18695	Human ORF ORF1628	c 176	17	0.9	529	21	AAI187203	3' end of cDNA seq
c 104	18	0.9	755	21	AAI18695	Arabidopsis thalia	c 177	17	0.9	529	21	AAI187203	Streptococcus miti
c 105	18	0.9	921	21	AAI18695	Arabidopsis thalia	c 178	17	0.9	560	22	AAI187203	Streptococcus miti
c 106	18	0.9	932	20	AAI18695	Arabidopsis thalia	c 179	17	0.9	560	22	AAI187203	Human digestive sy
c 107	18	0.9	1078	21	AAI18695	Arabidopsis thalia	c 180	17	0.9	561	22	AAI187203	Human digestive sy
c 108	18	0.9	1104	23	AAI18695	DNA encoding novel	c 181	17	0.9	574	23	AAI187203	Drosophila melanog
c 109	18	0.9	1197	21	AAI18695	Arabidopsis thalia	c 182	17	0.9	636	21	AAI187203	Human ATP binding
c 110	18	0.9	1200	22	AAI18695	Enterococcus faeca	c 183	17	0.9	639	21	AAI187203	Fusarium venenatu
c 111	18	0.9	1229	22	AAI18695	Human colon cancer	c 184	17	0.9	706	22	AAI187203	Mammalian vestibul
c 112	18	0.9	1281	21	AAI18695	Arabidopsis thalia	c 185	17	0.9	783	11	AAI187203	Fowl Pox Virus pro
c 113	18	0.9	1299	23	AAI18695	E. coli DNA for ce	c 186	17	0.9	794	22	AAI187203	Human neuroblastom
c 114	18	0.9	1319	21	AAI18695	Arabidopsis thalia	c 187	17	0.9	813	22	AAI187203	Human cDNA clone (
c 115	18	0.9	1348	21	AAI18695	Arabidopsis thalia	c 188	17	0.9	834	19	AAI187203	3' fragment of pro
c 116	18	0.9	1429	16	AAI18695	Neurotrophin-6 cDN	c 189	17	0.9	834	19	AAI187203	Human immunogenic
c 117	18	0.9	1564	22	AAI18695	Human purified sec	c 190	17	0.9	834	22	AAI187203	Human prostate cDN
c 118	18	0.9	1592	21	AAI18695	Human ORF ORF2605	c 191	17	0.9	834	22	AAI187203	Human prostate tum
c 119	18	0.9	1772	22	AAI18695	Human MIT-like pro	c 192	17	0.9	834	22	AAI187203	Human prostate-spe
c 120	18	0.9	1788	23	AAI18695	Drosophila melanog	c 193	17	0.9	834	22	AAI187203	Human prostate-tum
c 121	18	0.9	1975	22	AAI18695	Human cDNA sequenc	c 194	17	0.9	971	17	AAI187203	Human clone 41 gen
c 122	18	0.9	2064	22	AAI18695	Human polynucleoti	c 195	17	0.9	1042	22	AAI187203	Human digestive sy
c 123	18	0.9	2067	20	AAI18695	DNA methyltransfer	c 196	17	0.9	1043	22	AAI187203	Human digestive sy
c 124	18	0.9	2093	21	AAI18695	Human adult brain	c 197	17	0.9	1084	19	AAI187203	Human digestive sy
c 125	18	0.9	2288	21	AAI18695	Drosophila melanog	c 198	17	0.9	1084	19	AAI187203	Streptococcus pneu
c 126	18	0.9	2387	22	AAI18695	Human secreted pro	c 199	17	0.9	1095	23	AAI187203	DNA encoding novel
c 127	18	0.9	2739	23	AAI18695	Human polynucleoti	c 200	17	0.9	1134	23	AAI187203	Drosophila melanog
c 128	18	0.9	3327	23	AAI18695	DNA encoding novel	c 201	17	0.9	1202	19	AAI187203	H. pylori GHPO 143
c 129	18	0.9	3485	14	AAI18695	Drosophila melanog	c 202	17	0.9	1260	22	AAI187203	Escherichia coli p
c 130	18	0.9	3824	23	AAI18695	BAMW coat protein	c 203	17	0.9	1457	22	AAI187203	Human Nck-2 cDNA
c 131	18	0.9	4101	23	AAI18695	DNA encoding novel	c 204	17	0.9	1460	20	AAI187203	Human secreted pro
c 132	18	0.9	4994	23	AAI18695	Drosophila melanog	c 205	17	0.9	1498	22	AAI187203	Nucleotide sequenc
c 133	18	0.9	6250	23	AAI18695	DNA encoding novel	c 206	17	0.9	1531	21	AAI187203	Human secreted pro
c 134	18	0.9	6934	23	AAI18695	Drosophila melanog	c 207	17	0.9	1713	21	AAI187203	Arabidopsis thalia
c 135	18	0.9	7067	24	AAI18695	Long terminal repe	c 208	17	0.9	1714	21	AAI187203	Arabidopsis thalia
c 136	18	0.9	7783	23	AAI18695	DNA encoding novel	c 209	17	0.9	1719	23	AAI187203	DNA encoding novel
c 137	18	0.9	9359	23	AAI18695	DNA encoding novel	c 210	17	0.9	1812	22	AAI187203	Mouse IP3 sponge p
c 138	18	0.9	9767	23	AAI18695	Drosophila melanog	c 211	17	0.9	1813	22	AAI187203	Human cDNA sequenc
c 139	18	0.9	10280	22	AAI18695	Human immune/haema	c 212	17	0.9	1906	19	AAI187203	CD30 ligand gene u
c 140	18	0.9	10339	22	AAI18695	DNA encoding human	c 213	17	0.9	1984	23	AAI187203	Drosophila melanog
c 141	18	0.9	10339	22	AAI18695	Human reproductive	c 214	17	0.9	1985	20	AAI187203	Arabidopsis thalia
c 142	18	0.9	13165	22	AAI18695	Human pancreatic c	c 215	17	0.9	2013	22	AAI187203	Human colon cancer
c 143	18	0.9	13165	22	AAI18695	Human digestive sy	c 216	17	0.9	2258	23	AAI187203	Drosophila melanog
c 144	18	0.9	15186	22	AAI18695	Human nervous syst	c 217	17	0.9	2267	20	AAI187203	Human KDR signal t
c 145	18	0.9	15187	22	AAI18695	Human nervous syst	c 218	17	0.9	2347	22	AAI187203	Human cDNA sequenc
c 146	18	0.9	17758	22	AAI18695	Human immune/haema	c 219	17	0.9	2421	21	AAI187203	Arabidopsis thalia
c 147	18	0.9	19440	20	AAI18695	DNA methyltransfer	c 220	17	0.9	2429	23	AAI187203	Drosophila melanog
c 148	18	0.9	22008	22	AAI18695	Human nervous syst	c 221	17	0.9	2520	19	AAI187203	MSENO40-2 promote
c 149	18	0.9	22008	22	AAI18695	Human nervous syst	c 222	17	0.9	2594	23	AAI187203	Drosophila melanog
c 150	18	0.9	28690	20	AAI18695	Human cardiovascular	c 223	17	0.9	2672	24	AAI187203	Human prostate spe
c 151	18	0.9	39353	22	AAI18695	Enterococcus faeca	c 224	17	0.9	2710	21	AAI187203	Group B Streptococ
c 152	18	0.9	39353	22	AAI18695	Human immune/haema	c 225	17	0.9	2733	22	AAI187203	Human cDNA sequenc
c 153	18	0.9	39358	22	AAI18695	Human immune/haema	c 226	17	0.9	2741	22	AAI187203	Human cDNA sequenc
c 154	18	0.9	39358	22	AAI18695	Human immune/haema	c 227	17	0.9	2741	22	AAI187203	Human cDNA sequenc
c 155	18	0.9	49999	20	AAI18695	Human immune/haema	c 228	17	0.9	3435	22	AAI187203	S. epidermidis gen
										3484	19	AAI187203	Arabidopsis nemato
										3629	23	AAI187203	Drosophila melanog

229	17	0.9	3982	23	ABL03364	Drosophila melanog	c 302	16	0.8	26	24	AAD21908	PCR primer, 2767T
230	17	0.9	4032	22	ABL08104	Drosophila melanog	c 303	16	0.8	30	14	AAQ07067	Rat FC5R primer R2
231	17	0.9	4046	22	ABAB8881	Escherichia coli p	c 304	16	0.8	51	22	AAI77097	Human silent SNP c
232	17	0.9	4068	22	AAI60291	Human polynucleoti	c 305	16	0.8	51	22	RAH90159	Human interleukin-
233	17	0.9	4103	22	AAI58505	Human polynucleoti	c 306	16	0.8	133	21	AAH31079	Human secreted pro
234	17	0.9	4261	23	ABL24616	Drosophila melanog	307	16	0.8	140	19	AAI10526	Human biallelic po
235	17	0.9	4765	20	AAI13187	Enterococcus faeca	308	16	0.8	169	21	AAC24235	Human secreted pro
236	17	0.9	4793	23	ABL18712	Drosophila melanog	309	16	0.8	174	22	ABA66272	Human foetal liver
237	17	0.9	4794	23	ABL01874	Drosophila melanog	310	16	0.8	174	22	ABA33338	Probe #11804 for g
238	17	0.9	4805	23	ABL24621	Drosophila melanog	311	16	0.8	174	22	RAK14689	Human brain expres
239	17	0.9	4950	23	ABL21378	Drosophila melanog	312	16	0.8	174	22	RAK40433	Human bone marrow
240	17	0.9	5177	22	AAH56333	Rat soluble adenyli	c 313	16	0.8	175	21	AAC24542	Human secreted pro
241	17	0.9	5177	22	AAH77294	Polynucleotide of	c 314	16	0.8	186	22	AAK19233	Human brain expres
242	17	0.9	5675	23	ABL16334	Drosophila melanog	c 315	16	0.8	186	22	AAK45203	Human bone marrow
243	17	0.9	5820	23	ABL03878	Drosophila melanog	c 316	16	0.8	217	21	AAC22600	Human secreted pro
244	17	0.9	6133	24	AAH61052	Human gene regulat	c 317	16	0.8	227	16	AAK24337	Human gene signatu
245	17	0.9	6287	23	ABL03876	Drosophila melanog	c 318	16	0.8	248	22	AAK58739	Human immune/haema
246	17	0.9	6623	19	AAV19051	Mitochondrial gene	c 319	16	0.8	256	20	AAH06693	Human immune/haema
247	17	0.9	6623	19	AAV19053	Mitochondrial gene	320	16	0.8	256	20	AAH06693	Chromosome 8 deriv
248	17	0.9	7148	23	ABL22002	Drosophila melanog	c 321	16	0.8	258	21	AAA61124	SEN virus genome f
249	17	0.9	7191	23	ABL14461	Drosophila melanog	c 322	16	0.8	264	21	AAC20818	Human secreted pro
250	17	0.9	7574	22	AAK74485	Drosophila melanog	c 323	16	0.8	285	21	AAK18897	Human secreted pro
251	17	0.9	7574	22	AAK74485	Human immune/haema	c 324	16	0.8	288	14	AAQ61137	Human brain Expres
252	17	0.9	8414	22	AAK74613	Human immune/haema	c 325	16	0.8	289	22	AAK58095	Human immune/haema
253	17	0.9	9388	23	ABL24620	DNA encoding novel	c 326	16	0.8	292	17	AAI34605	Probe for detectin
254	17	0.9	9848	21	AAZ98272	Drosophila melanog	327	16	0.8	300	20	AAI21794	Human gene express
255	17	0.9	9871	12	AAQ13593	Mouse IP3 receptor	328	16	0.8	306	21	AAC50168	Arabidopsis thalia
256	17	0.9	10006	24	ABL32036	Human immune syste	c 329	16	0.8	303	21	AAH70484	HIV Tat_Cys22_SF16
257	17	0.9	10317	23	ABL14460	Human immune/haema	c 330	16	0.8	310	22	AAK83981	Human immune/haema
258	17	0.9	12714	23	ABL14976	Drosophila melanog	c 331	16	0.8	327	21	AAC45039	Arabidopsis thalia
259	17	0.9	12893	21	AAAB1733	N. meningitidis pa	c 332	16	0.8	332	22	AAK63065	Human immune/haema
260	17	0.9	13982	22	AAK65456	Human immune/haema	c 333	16	0.8	345	22	AAK58749	Human immune/haema
261	17	0.9	13982	22	AAK77762	Human immune/haema	c 334	16	0.8	348	20	AAV86174	EST clone J1168. H
262	17	0.9	13982	22	AAK84022	Human immune/haema	c 335	16	0.8	349	14	AAQ61138	Human brain Expres
263	17	0.9	16287	24	ABL32672	Human immune syste	c 336	16	0.8	353	15	AAQ44781	Drosophila anti-ba
264	17	0.9	17131	21	AAZ60888	DNA encoding a hum	c 337	16	0.8	357	23	AAH52011	Staphylococcus aur
265	17	0.9	18920	23	ABL02306	Drosophila melanog	c 338	16	0.8	359	18	AAI75248	Nucleotide sequenc
266	17	0.9	21126	20	AAK12960	Enterococcus faeca	c 339	16	0.8	363	15	AAQ44782	Drosophila anti-ba
267	17	0.9	22680	22	AAK66308	Human immune/haema	c 340	16	0.8	363	21	AAC36501	Human secreted pro
268	17	0.9	22680	22	AAK73334	Human immune/haema	c 341	16	0.8	363	21	AAI11914	Human breast cance
269	17	0.9	22680	22	AAK73344	Human immune/haema	c 342	16	0.8	363	22	AAK59926	Human immune/haema
270	17	0.9	22680	22	AAK73625	Human immune/haema	c 343	16	0.8	364	22	AAI20804	Human breast cance
271	17	0.9	22680	22	AAK73847	Human immune/haema	c 344	16	0.8	380	22	AAH22792	Human cDNA encodin
272	17	0.9	22680	22	AAK73934	Human immune/haema	c 345	16	0.8	388	22	AAI09550	Human breast cance
273	17	0.9	22680	22	AAK78350	Human immune/haema	c 346	16	0.8	389	22	AAI22558	Human breast cance
274	17	0.9	22788	23	ABL20208	Drosophila melanog	c 347	16	0.8	390	22	AAK74773	Human immune/haema
275	17	0.9	24908	22	AAK81665	Human immune/haema	c 348	16	0.8	392	21	AAC07428	Human secreted pro
276	17	0.9	29973	22	AAH39814	Human cytoskeletal	c 349	16	0.8	393	22	ABA11994	Human nervous syst
277	17	0.9	31034	23	ABL13678	Drosophila melanog	c 350	16	0.8	400	18	AAV76468	Staphylococcus aur
278	17	0.9	47475	21	AAAB1465	N. meningitidis pa	c 351	16	0.8	402	22	AAH22302	Human immunoglobul
279	17	0.9	48203	22	AAK70161	Human immune/haema	c 352	16	0.8	406	22	AAH29374	Drosophila melanog
280	17	0.9	48203	22	AAK81663	Human immune/haema	c 353	16	0.8	417	21	AAH61158	SEN virus genome f
281	17	0.9	48203	22	AAK82628	Human immune/haema	c 354	16	0.8	421	22	AAI13689	Human breast cance
282	17	0.9	48204	22	AAK70164	Human immune/haema	c 355	16	0.8	423	20	AAH88755	EST clone HK669.
283	17	0.9	48204	22	AAK81666	Human immune/haema	c 356	16	0.8	423	21	AAH16425	Human prostate can
284	17	0.9	48204	22	AAK82630	Human immune/haema	c 357	16	0.8	432	21	AAK60625	Human immune/haema
285	17	0.9	49431	23	ABL13742	Drosophila melanog	c 358	16	0.8	432	21	AAC26669	Human secreted pro
286	17	0.9	49767	21	AAAB1458	N. meningitidis pa	c 359	16	0.8	435	22	AAK76539	Human immune/haema
287	17	0.9	54863	22	AAK86025	Human immune/haema	c 360	16	0.8	438	24	ABK16439	Gram positive bact
288	17	0.9	54877	22	AAK86026	Human immune/haema	c 361	16	0.8	438	24	ABK16464	Gram positive bact
289	17	0.9	58407	19	AAV21210	Methanococcus jann	c 362	16	0.8	449	22	ABA46193	Human breast cell
290	17	0.9	111309	20	AAK20250	Borrelia burgdorfe	c 363	16	0.8	449	22	ABA56733	Human foetal liver
291	17	0.9	151826	21	AAK22291	BAC containing rep	c 364	16	0.8	449	22	ABA26359	Probe #4825 for ge
292	17	0.9	172325	21	AAK21613	Neisseria meningit	c 365	16	0.8	449	22	AAK04865	Human brain expres
293	17	0.9	349980	21	AAK21608	Neisseria meningit	c 366	16	0.8	449	22	AAK30393	Human bone marrow
294	17	0.9	349980	21	AAK21612	Neisseria meningit	c 367	16	0.8	449	22	AAI15001	Probe #4934 for ge
295	17	0.9	349980	22	AAH1226	Pyrococcus abyssi	c 368	16	0.8	449	22	AAI36346	Probe #5032 used t
296	17	0.9	837096	21	AAAB1489	N. meningitidis pa	c 369	16	0.8	449	22	AAI04766	Probe #4757 used t
297	17	0.9	1038602	20	AAZ01425	Complete genome se	c 370	16	0.8	451	21	AAC24241	Human secreted pro
298	17	0.9	1664976	19	AAV21209	Methanococcus jann	c 371	16	0.8	456	23	AAH54920	Staphylococcus aur
299	16	0.8	20	22	AAH26600	Human Nck-2 phosph	c 372	16	0.8	456	11	AAQ05318	Sequence encoding
300	16	0.8	21	22	AAH26600	Human Mda-7 gene p	c 373	16	0.8	458	12	AAQ11644	Sequence encoding
301	16	0.8	24	24	ABK16600	CXCR2 receptor PCR	c 374	16	0.8	465	22	AAK06466	Human brain expres

c 375	16	0.8	465	22	AAK32147	Human bone marrow	448	16	0.8	700	22	AAH92519	Human inflammatory
376	16	0.8	470	22	ABA53690	Human foetal liver	c 449	16	0.8	700	22	AAH93259	Human inflammatory
377	16	0.8	470	22	ABA23443	Probe #1909 for ge	450	16	0.8	707	24	AAH93259	Human G protein-co
378	16	0.8	470	22	AAK01954	Human brain expres	c 451	16	0.8	714	22	AAH93259	H. pylori HPC005 e
379	16	0.8	470	22	AAK27412	Human bone marrow	c 452	16	0.8	722	22	AAH93259	Human CDNA clone (
c 380	16	0.8	471	22	ABA43971	Human breast cell	c 453	16	0.8	724	21	AAH93259	Aspergillus oryzae
c 381	16	0.8	471	22	ABA54424	Human foetal liver	454	16	0.8	732	22	AAH93259	Human neuroblastom
c 382	16	0.8	471	22	ABA24206	Probe #2672 for ge	455	16	0.8	732	22	AAH93259	Human prostate can
c 383	16	0.8	471	22	AAH08088	Human breast thallia	c 456	16	0.8	747	22	AAH93259	Human musculoskele
c 384	16	0.8	473	21	AAK37200	Arabidopsis thallia	c 457	16	0.8	748	22	AAH93259	Human neuroblastom
c 385	16	0.8	475	22	ABA59103	Human foetal liver	c 458	16	0.8	766	19	AAH93259	DNA encoding a S.
c 386	16	0.8	475	22	AAK07282	Human brain expres	c 459	16	0.8	767	22	AAH93259	Human CDNA encodin
c 387	16	0.8	475	22	AAK33042	Human bone marrow	c 460	16	0.8	773	19	AAH93259	Human mu-opioid re
c 388	16	0.8	475	22	AAI38845	Probe #7531 used t	c 461	16	0.8	773	19	AAH93259	Human mu-opioid re
c 389	16	0.8	481	22	AAH01692	Human reproductive	c 462	16	0.8	773	19	AAH93259	Human mu-opioid re
c 390	16	0.8	481	22	AAH26763	Human genomic DNA	c 463	16	0.8	773	19	AAH93259	Human mu-opioid re
c 391	16	0.8	481	22	AAH26764	Human genomic DNA	c 464	16	0.8	773	19	AAH93259	Human mu-opioid re
c 392	16	0.8	486	20	AAH84692	Human metastatic m	c 465	16	0.8	773	19	AAH93259	Human mu-opioid re
c 393	16	0.8	495	22	AAH18567	Streptococcus pneu	c 466	16	0.8	773	19	AAH93259	Human mu-opioid re
c 394	16	0.8	495	22	AAH26264	Human CDNA encodin	c 467	16	0.8	773	19	AAH93259	Human mu-opioid re
c 395	16	0.8	505	22	ABA19800	Human nervous syst	c 468	16	0.8	786	20	AAH93259	Human validated ca
c 396	16	0.8	505	22	ABA19801	Human nervous syst	c 469	16	0.8	791	22	AAH93259	Human CDNA clone (
c 397	16	0.8	510	22	AAH92639	Human inflammatory	c 470	16	0.8	799	21	AAH93259	Arabidopsis thallia
c 398	16	0.8	517	22	AAH12278	Human cytochrome p	c 471	16	0.8	801	22	AAH93259	Human CDNA 5'-end
c 399	16	0.8	526	21	AAH45585	Arabidopsis thallia	c 472	16	0.8	810	22	AAH93259	Streptococcus pneu
c 400	16	0.8	528	22	AAH68562	Human immune/haema	c 473	16	0.8	816	22	AAH93259	Human CDNA clone r
c 401	16	0.8	528	22	AAH68563	Human immune/haema	c 474	16	0.8	816	22	AAH93259	Human CDNA clone (
c 402	16	0.8	534	22	AAH14944	Human breast cance	c 475	16	0.8	826	23	AAH93259	DNA encoding novel
c 403	16	0.8	534	22	AAH84812	Human immune/haema	c 476	16	0.8	831	21	AAH93259	Streptococcus pneu
c 404	16	0.8	539	21	AAH97996	Human colon cancer	c 477	16	0.8	831	21	AAH93259	CFE 36 coding sequ
c 405	16	0.8	541	21	AAH41938	Arabidopsis thallia	c 478	16	0.8	834	22	AAH93259	Human immune/haema
c 406	16	0.8	549	21	AAH12602	Aspergillus oryzae	c 479	16	0.8	834	22	AAH93259	Ehrlichia canis nu
c 407	16	0.8	549	21	AAH94795	Cat flea hindgut p	c 480	16	0.8	849	21	AAH93259	Trichoderma reesei
c 408	16	0.8	550	22	AAH12277	Human cytochrome p	c 481	16	0.8	854	22	AAH93259	2CFE 36 coding seq
c 409	16	0.8	570	21	AAH50166	Arabidopsis thallia	c 482	16	0.8	860	21	AAH93259	Arabidopsis thallia
c 410	16	0.8	574	19	AAH60062	Nucleic acid GII5c	c 483	16	0.8	865	21	AAH93259	Arabidopsis thallia
c 411	16	0.8	578	22	AAH16635	Human nervous syst	c 484	16	0.8	868	22	AAH93259	Human CDNA clone (
c 412	16	0.8	578	22	AAH35605	Human cardiovascular	c 485	16	0.8	875	18	AAH93259	DNA encoding a Sta
c 413	16	0.8	579	22	AAH02068	Streptococcus pneu	c 486	16	0.8	875	19	AAH93259	DNA encoding a Sta
c 414	16	0.8	586	21	AAH08837	Fusarium venenatum	c 487	16	0.8	880	21	AAH93259	Arabidopsis thallia
c 415	16	0.8	590	21	AAH33873	Arabidopsis thallia	c 488	16	0.8	881	21	AAH93259	Mouse immortalizat
c 416	16	0.8	592	22	AAH23079	Osteoarthritis tis	c 489	16	0.8	890	19	AAH93259	Nucleic acid G256
c 417	16	0.8	594	22	AAH35599	Human foetal liver	c 490	16	0.8	891	21	AAH93259	Arabidopsis thallia
c 418	16	0.8	594	22	AAH12125	Human brain expres	c 491	16	0.8	891	21	AAH93259	Arabidopsis thallia
c 419	16	0.8	594	22	AAH37842	Human bone marrow	c 492	16	0.8	904	21	AAH93259	Arabidopsis thallia
c 420	16	0.8	594	22	AAH43715	Probe #12401 used	c 493	16	0.8	930	23	AAH93259	Salmonella typhi D
c 421	16	0.8	595	19	AAH47654	Nucleotide sequenc	c 494	16	0.8	951	22	AAH93259	Human olfactory re
c 422	16	0.8	601	21	AAH40754	Arabidopsis thallia	c 495	16	0.8	959	20	AAH93259	Melanoma kinase DN
c 423	16	0.8	601	22	AAH28979	Drosophila melanog	c 496	16	0.8	960	23	AAH93259	DNA encoding novel
c 424	16	0.8	607	22	AAH46903	Human G protein-co	c 497	16	0.8	978	21	AAH93259	Human UGT2B15 exon
c 425	16	0.8	609	21	AAH77030	Human ORFX ORF2585	c 498	16	0.8	985	24	AAH93259	Human immune syste
c 426	16	0.8	609	21	AAH46165	Arabidopsis thallia	c 499	16	0.8	988	20	AAH93259	Enterococcus faeca
c 427	16	0.8	611	22	AAH23800	Human breast cance	c 500	16	0.8	1029	23	AAH93259	Haemophilus influe
c 428	16	0.8	618	22	AAH11628	Human foetal liver	c 501	16	0.8	1041	22	AAH93259	Escherichia coli p
c 429	16	0.8	618	22	AAH19966	Human brain expres	c 502	16	0.8	1041	23	AAH93259	E. coli DNA for ce
c 430	16	0.8	618	22	AAH46003	Human bone marrow	c 503	16	0.8	1054	20	AAH93259	Human secreted pro
c 431	16	0.8	618	22	AAH51918	Probe #20604 used	c 504	16	0.8	1055	21	AAH93259	Arabidopsis thallia
c 432	16	0.8	622	20	AAH87425	EST clone BR559.	c 505	16	0.8	1055	22	AAH93259	Human immune/haema
c 433	16	0.8	625	21	AAH08607	Fusarium venenatum	c 506	16	0.8	1059	21	AAH93259	Arabidopsis thallia
c 434	16	0.8	643	19	AAH98650	DNA encoding a S.	c 507	16	0.8	1062	21	AAH93259	Arabidopsis thallia
c 435	16	0.8	645	22	AAH26765	Human genomic DNA	c 508	16	0.8	1078	22	AAH93259	Human immune/haema
c 436	16	0.8	646	22	AAH26767	Human genomic DNA	c 509	16	0.8	1092	20	AAH93259	Enterococcus faeca
c 437	16	0.8	654	22	AAH34120	Human CDNA encodin	c 510	16	0.8	1112	18	AAH93259	DNA encoding a S.
c 438	16	0.8	673	21	AAH13747	Aspergillus oryzae	c 511	16	0.8	1146	21	AAH93259	Group B Streptococ
c 439	16	0.8	673	23	AAH83171	DNA encoding novel	c 512	16	0.8	1155	22	AAH93259	Arabidopsis thallia
c 440	16	0.8	676	21	AAH47722	Arabidopsis thallia	c 513	16	0.8	1161	22	AAH93259	Drosophila gustato
c 441	16	0.8	679	22	ABA11164	Human nervous syst	c 514	16	0.8	1185	21	AAH93259	Arabidopsis thallia
c 442	16	0.8	682	22	AAH03483	Human CDNA clone (c 515	16	0.8	1186	13	AAH93259	Rubisco promoted h
c 443	16	0.8	683	22	AAH21664	Human breast cance	c 516	16	0.8	1194	22	AAH93259	Human AFP protein
c 444	16	0.8	697	21	AAH13471	Aspergillus oryzae	c 517	16	0.8	1232	17	AAH93259	Probe for detectin
c 445	16	0.8	699	22	AAH92104	Human inflammatory	c 518	16	0.8	1257	23	AAH93259	DNA encoding novel
c 446	16	0.8	700	22	AAH92105	Human inflammatory	c 519	16	0.8	1262	21	AAH93259	Human secreted pro
c 447	16	0.8	700	22	AAH92518	Human inflammatory	c 520	16	0.8	1279	22	AAH93259	Human polynucleoti
										1287	22	ABA08873	Human secreted pro

c 521	16	0.8	1295	19	AAV11659	Rat spleen ICAM-4	c 594	16	0.8	1721	11	AAQ06319	Synthetic gene enc
c 522	16	0.8	1295	22	AAV19330	Rat spleen partial	c 595	16	0.8	1721	12	AAQ13164	Gene encoding t-PA
c 523	16	0.8	1295	22	AAV19330	Partial rat ICAM-R	c 596	16	0.8	1721	12	AAQ13165	Gene encoding t-PA
c 524	16	0.8	1314	18	AAV63436	Mangosteen Class I	c 597	16	0.8	1721	20	AAH07566	Homo sapiens fetal
c 525	16	0.8	1314	18	AAV63436	Mangosteen Class I	c 598	16	0.8	1722	20	AAH65626	C glutamicum codin
c 526	16	0.8	1314	20	AAV71847	Mangosteen C18:1 a	c 599	16	0.8	1727	14	AAQ46262	Encodes modified t
c 527	16	0.8	1317	22	AAV29058	cdNA encoding for	c 600	16	0.8	1732	23	AAH77812	DNA encoding novel
c 528	16	0.8	1353	17	AAV41706	Lymphocyte specifi	c 601	16	0.8	1738	22	AAH35761	Human musculoskele
c 529	16	0.8	1353	21	AAZ39862	Plasmodium ovale p	c 602	16	0.8	1739	22	AAH35760	Human musculoskele
c 530	16	0.8	1387	22	AAZ39862	cdNA encoding for	c 603	16	0.8	1743	20	AAV72028	Adenovirus wild-ty
c 531	16	0.8	1400	15	AAQ70142	Human cartilage li	c 604	16	0.8	1743	21	AAH37052	Human PRO1315 (UNQ
c 532	16	0.8	1420	22	AAV06495	Arabidopsis thalia	c 605	16	0.8	1743	22	AAH92095	Human PRO1315 cDNA
c 533	16	0.8	1435	19	AAV16466	Nucleotide sequenc	c 606	16	0.8	1743	22	AAH54274	DNA encoding prote
c 534	16	0.8	1449	23	ABL26303	Drosophila melanog	c 607	16	0.8	1746	21	AAH01223	Adenovirus 5 (Ad5)
c 535	16	0.8	1456	24	AAH594703	Rat secreted facto	c 608	16	0.8	1746	22	AAH31329	Recombinant adenov
c 536	16	0.8	1472	19	AAV34660	Rodent cDNA clone	c 609	16	0.8	1748	16	AAQ099008	Interleukin 8 rece
c 537	16	0.8	1472	19	AAV11662	Rat ICAM-4 clone E	c 610	16	0.8	1749	22	AAH17031	Human cDNA sequenc
c 538	16	0.8	1472	19	AAV19333	Rat ICAM-4 partial	c 611	16	0.8	1750	21	AAH20991	Human low adenosin
c 539	16	0.8	1472	20	AAH36492	Rat ICAM coding se	c 612	16	0.8	1750	21	AAH34869	Human adenosine re
c 540	16	0.8	1472	20	AAV08979	Rat ICAM-4 coding	c 613	16	0.8	1751	21	AAH27133	Human inflammation
c 541	16	0.8	1474	21	AAH58737	Arabidopsis thalia	c 614	16	0.8	1755	14	AAQ34490	Insert in pCGP175
c 542	16	0.8	1479	23	ABL19975	Human polynucleoti	c 615	16	0.8	1759	11	AAQ01277	DNA fragment carry
c 543	16	0.8	1481	13	AAQ23496	Drosophila melanog	c 616	16	0.8	1801	18	AAH67080	Mouse oligodendroc
c 544	16	0.8	1486	22	AAH65954	Encodes pre t-PA v	c 617	16	0.8	1801	22	AAH89708	Mouse oligodendroc
c 545	16	0.8	1486	22	AAH65954	Nucleotide sequenc	c 618	16	0.8	1801	24	AAH99895	Mouse ischaemic co
c 546	16	0.8	1497	22	AAH60467	Human polynucleoti	c 619	16	0.8	1807	21	AAH20992	Human low adenosin
c 547	16	0.8	1502	22	AAH79031	Human immune/haema	c 620	16	0.8	1807	21	AAH34870	Human adenosine re
c 548	16	0.8	1502	22	AAH79033	Human immune/haema	c 621	16	0.8	1812	14	AAQ34489	Insert in pCGP176
c 549	16	0.8	1510	14	AAH38747	Human IL-8 recepto	c 622	16	0.8	1812	14	AAQ49417	Cytochrome P450 ho
c 550	16	0.8	1510	21	AAH20993	Human low adenosin	c 623	16	0.8	1818	21	AAH93819	Modified fibre pro
c 551	16	0.8	1510	21	AAH34871	Human adenosine re	c 624	16	0.8	1818	21	AAH93821	Modified fibre pro
c 552	16	0.8	1512	23	ABL22001	Drosophila melanog	c 625	16	0.8	1824	14	AAQ47842	Flavonoid-3',5'-ny
c 553	16	0.8	1515	19	AAV52022	Helicobacter poly	c 626	16	0.8	1837	21	AAH34806	Arabidopsis thalia
c 554	16	0.8	1524	22	AAH88103	Human FLEHT-34 nu	c 627	16	0.8	1848	21	AAH93820	Modified fibre pro
c 555	16	0.8	1528	13	AAQ28715	Wild-type AOX2 pro	c 628	16	0.8	1848	21	AAH93822	Modified fibre pro
c 556	16	0.8	1528	13	AAQ28717	Mutant AOX2 promot	c 629	16	0.8	1884	21	AAH15985	Human prostate can
c 557	16	0.8	1528	15	AAQ63285	Natural AOX2 promo	c 630	16	0.8	1886	22	AAH09268	Human IL-8R B homo
c 558	16	0.8	1528	15	AAQ63289	Mutant AOX2 promot	c 631	16	0.8	1893	21	AAH93823	Modified fibre pro
c 559	16	0.8	1528	16	AAH081992	Alcohol oxidase (A	c 632	16	0.8	1893	21	AAH93824	Modified fibre pro
c 560	16	0.8	1529	22	AAH03388	Human polynucleoti	c 633	16	0.8	1920	23	AAH592476	DNA encoding novel
c 561	16	0.8	1530	22	AAH52352	S. epidermidis ope	c 634	16	0.8	1923	21	AAH93825	Modified fibre pro
c 562	16	0.8	1536	21	AAH34328	Arabidopsis thalia	c 635	16	0.8	1923	21	AAH93826	Modified fibre pro
c 563	16	0.8	1538	21	AAH47897	Arabidopsis thalia	c 636	16	0.8	1926	21	AAH61125	SEN virus genome f
c 564	16	0.8	1547	13	AAQ28718	Mutant AOX2 promot	c 637	16	0.8	1931	19	AAH03142	Aspergillus nidula
c 565	16	0.8	1547	15	AAQ63291	Mutant AOX2 promot	c 638	16	0.8	1931	20	AAH227413	A. nidulans phytas
c 566	16	0.8	1553	22	AAH58602	Human polynucleoti	c 639	16	0.8	1944	23	AAH55607	Streptococcus pneu
c 567	16	0.8	1565	21	AAH38033	Arabidopsis thalia	c 640	16	0.8	1963	22	AAH08826	Human class II cyt
c 568	16	0.8	1568	19	AAH68326	Human STIF cDNA	c 641	16	0.8	1973	24	AAH26494	Sheep alpha (1, 3)
c 569	16	0.8	1579	23	AAH86135	DNA encoding novel	c 642	16	0.8	1976	21	AAH95206	Human UDP-glucuron
c 570	16	0.8	1588	21	AAH29800	Sequence upstream	c 643	16	0.8	1983	22	AAH18571	Human cDNA sequenc
c 571	16	0.8	1595	21	AAH22417	Human secreted pro	c 644	16	0.8	2009	22	AAH05410	Human secreted pro
c 572	16	0.8	1605	11	AAQ06317	Synthetic gene enc	c 645	16	0.8	2018	21	AAH26865	Essential Staphylo
c 573	16	0.8	1605	13	AAQ23498	Encodes t-PA varia	c 646	16	0.8	2018	22	AAH08016	Staphylococcus aur
c 574	16	0.8	1605	14	AAQ46260	Synthetic tPA gene	c 647	16	0.8	2018	22	AAH91547	Staphylococcus aur
c 575	16	0.8	1609	10	AAH90742	Chemically synthe	c 648	16	0.8	2024	21	AAH94052	Haemophilus contort
c 576	16	0.8	1616	10	AAH90744	Synthetic gene cod	c 649	16	0.8	2039	22	AAH45110	DNA encoding nove
c 577	16	0.8	1620	22	AAH29145	Polynucleotide seq	c 650	16	0.8	2066	23	AAH81269	DNA encoding novel
c 578	16	0.8	1639	16	AAH74059	Human interleukin-	c 651	16	0.8	2075	19	AAH16467	Nucleotide sequenc
c 579	16	0.8	1639	20	AAH14999	DNA encoding an in	c 652	16	0.8	2076	22	AAH56326	DNA encoding Chlam
c 580	16	0.8	1639	22	AAH02920	Human shear stress	c 653	16	0.8	2079	21	AAH61156	SEN virus genome f
c 581	16	0.8	1652	21	AAH79011	Human secreted pro	c 654	16	0.8	2088	17	AAH36895	Candida albicans l
c 582	16	0.8	1660	21	AAH75849	Human ORFX ORF1404	c 655	16	0.8	2088	22	AAH07947	Human amyloid prec
c 583	16	0.8	1660	22	AAH08323	Human Cdc42-bindin	c 656	16	0.8	2091	20	AAH18839	Streptococcus equi
c 584	16	0.8	1673	24	AAH562329	cDNA sequence #116	c 657	16	0.8	2106	22	AAH94726	Human full-length
c 585	16	0.8	1686	22	AAH57487	Human liver cell s	c 658	16	0.8	2106	22	AAH94726	Human full-length
c 586	16	0.8	1689	23	AAH52961	DNA encoding novel	c 659	16	0.8	2114	23	AAH67980	DNA encoding novel
c 587	16	0.8	1700	19	AAH32893	Human interleukin-	c 660	16	0.8	2136	22	AAH60006	Human synthetase #
c 588	16	0.8	1700	20	AAH10787	Human tumour supre	c 661	16	0.8	2149	22	AAH99560	Human protein enco
c 589	16	0.8	1705	22	AAH46121	Human DNA encoding	c 662	16	0.8	2149	22	AAH99560	Human protein enco
c 590	16	0.8	1705	22	AAH44264	Human PRO3301 nucl	c 663	16	0.8	2161	21	AAH21788	Human breast and o
c 591	16	0.8	1706	21	AAH40491	Human fetal kidney	c 664	16	0.8	2176	23	ABL10759	Drosophila melanog
c 592	16	0.8	1718	21	AAH90052	Melanoma different	c 665	16	0.8	2184	23	ABL15025	Drosophila melanog
c 593	16	0.8	1719	21	AAH51095	Arabidopsis thalia	c 666	16	0.8	2195	23	ABL16866	Drosophila melanog

667	16	0.8	2197	22	ABA07287	Human pancreatic c	740	16	0.8	3373	23	ABL06164	Drosophila melanog
668	16	0.8	2197	22	AAK89932	Human digestive sy	c 741	16	0.8	3394	22	AAH54265	S. epidermidis gen
669	16	0.8	2214	18	AAAT84275	Rat ICAM-4 cDNA RT	c 742	16	0.8	3416	22	AAH81784	Human differential
670	16	0.8	2214	19	AAV34658	RT-PCR clone for c	743	16	0.8	3428	22	AAI07433	Human reproductive
671	16	0.8	2214	19	AAV11660	Rat ICAM-4 cDNA fr	744	16	0.8	3428	22	AAI07433	Human breast or ov
672	16	0.8	2214	19	AAV19331	Rat ICAM-4 RT-PCR	c 745	16	0.8	3476	23	ABLI19970	Drosophila melanog
673	16	0.8	2214	20	AAK36497	Rat ICAM coding se	c 746	16	0.8	3483	23	ABLI19970	Drosophila melanog
674	16	0.8	2214	20	AAV08977	Rat ICAM-4 coding	c 747	16	0.8	3516	19	AAV57515	zcytor7 cytokine r
675	16	0.8	2247	18	AAV25134	H. pylori inner me	748	16	0.8	3516	22	AAV57515	Human cDNA encodin
676	16	0.8	2247	18	AAV24920	H. pylori ORF 05a	749	16	0.8	3516	22	AAV57515	Human IL-20 recept
677	16	0.8	2260	22	AAV40661	DNA encoding human	750	16	0.8	3532	21	AAH22815	Mouse homologue of
678	16	0.8	2260	22	AAV40661	Human reproductive	751	16	0.8	3562	21	AAH22815	Human muscleoskele
679	16	0.8	2268	21	AAK06349	DNA encoding human	752	16	0.8	3573	23	ABLI12105	Drosophila melanog
680	16	0.8	2269	21	AAK39419	Human ets2 coding	753	16	0.8	3595	23	ABLI26302	Drosophila melanog
681	16	0.8	2286	22	AAH26595	Human melanoma dif	c 754	16	0.8	3598	22	AAV93803	Human cDNA encodin
682	16	0.8	2296	23	AAV77847	DNA encoding novel	c 755	16	0.8	3601	22	AAH54151	S. epidermidis gen
683	16	0.8	2314	21	AAK76745	Human ORFX ORF2300	756	16	0.8	3608	21	AAV99916	CDNA encoding huma
684	16	0.8	2329	19	AAK14596	H. pylori GHP0 175	757	16	0.8	3608	22	AAH46141	Human DNA encoding
685	16	0.8	2360	22	AAV07734	Human secreted pro	758	16	0.8	3635	17	AAV06761	Human c-mer protoo
686	16	0.8	2368	21	AAK26851	Essential Staphylo	759	16	0.8	3635	23	ABLI19872	Drosophila melanog
687	16	0.8	2368	22	AAV08002	Staphylococcus aur	760	16	0.8	3661	23	ABLI22000	Drosophila melanog
688	16	0.8	2368	22	AAV91533	Staphylococcus aur	761	16	0.8	3692	22	AAH33758	Human colon cancer
689	16	0.8	2373	20	AAK13400	Enterococcus faeca	c 762	16	0.8	3746	21	AAV99512	DNA encoding a mai
690	16	0.8	2396	23	ABLI25394	Drosophila melanog	c 763	16	0.8	3746	21	AAV99512	DNA encoding a mai
691	16	0.8	2400	23	AAV51712	Staphylococcus aur	c 764	16	0.8	3749	23	ABLI19974	Drosophila melanog
692	16	0.8	2403	17	AAV12562	S.aureus topoisome	c 765	16	0.8	3773	21	AAV99494	DNA encoding a mai
693	16	0.8	2403	17	AAV12573	S.aureus mutant gr	c 766	16	0.8	3799	24	AAV16458	Corn cDNA encoding
694	16	0.8	2403	22	AAH01330	Staphylococcus aur	c 767	16	0.8	3801	22	AAV05465	Human reproductive
695	16	0.8	2403	23	AAV54404	Staphylococcus aur	c 768	16	0.8	3866	22	AAH54616	S. epidermidis gen
696	16	0.8	2403	23	AAV55401	Staphylococcus aur	c 769	16	0.8	3881	21	AAV61269	Human secreted pro
697	16	0.8	2405	22	AAV02351	Human RNA metaboli	c 770	16	0.8	3891	21	AAV61269	Human ORFX ORF1979
698	16	0.8	2469	21	AAV65740	Streptococcus pneu	c 771	16	0.8	3935	21	AAV98913	Human pancreatic c
699	16	0.8	2469	21	AAV05811	Group B Streptococ	772	16	0.8	3955	23	ABLI25064	Drosophila melanog
700	16	0.8	2469	22	AAV00038	Streptococcus agal	773	16	0.8	3957	23	AAV78289	DNA encoding novel
701	16	0.8	2472	21	AAV65741	Streptococcus pneu	774	16	0.8	4080	23	ABLI03766	Drosophila melanog
702	16	0.8	2478	22	AAV00036	Streptococcus pyog	c 775	16	0.8	4099	23	AAV86847	DNA encoding novel
703	16	0.8	2480	22	AAV15977	Human cDNA sequenc	c 776	16	0.8	4145	24	ABV99868	Mouse ischaemic co
704	16	0.8	2550	19	AAV34661	Rat brain cDNA clo	c 777	16	0.8	4223	22	AAV158965	Human polynucleoti
705	16	0.8	2550	19	AAV11663	Rat ICAM-4 clone 7	c 778	16	0.8	4248	21	AAV160751	Human polynucleoti
706	16	0.8	2550	19	AAV19334	Rat brain ICAM-4 p	779	16	0.8	4254	23	ABLI13025	S. aureus_RUSA266
707	16	0.8	2550	20	AAV08980	Rat ICAM-4 coding	780	16	0.8	4301	23	ABLI12888	Drosophila melanog
708	16	0.8	2567	21	AAV35376	Absorptive hyperca	781	16	0.8	4305	20	AAV87945	Candida claoace fa
709	16	0.8	2568	22	AAV15184	Human cDNA sequenc	c 782	16	0.8	4373	18	AAV72719	Pyruvate:flavodoxi
710	16	0.8	2579	22	AAV60794	Human polynucleoti	783	16	0.8	4394	22	AAV35762	Human muscleoskele
711	16	0.8	2646	21	AAV61122	SEN virus genome f	c 784	16	0.8	4400	22	AAV35763	Human muscleoskele
712	16	0.8	2665	23	AAV74027	DNA encoding novel	c 785	16	0.8	4468	22	AAV158364	Human polynucleoti
713	16	0.8	2760	21	AAV51152	Arabidopsis thalia	c 786	16	0.8	4495	17	AAV32143	Helicobacter-speci
714	16	0.8	2760	22	AAV07710	Human secreted pro	c 787	16	0.8	4505	16	AAQ88164	Tomato TGTR1 ethy
715	16	0.8	2780	20	AAV61239	Soybean raffinose	c 788	16	0.8	4566	19	AAV59048	Tomato ethylene re
716	16	0.8	2784	22	AAV18856	Human nervous syst	c 789	16	0.8	4590	22	AAV18208	Human cDNA sequenc
717	16	0.8	2785	23	AAV77796	DNA encoding novel	c 790	16	0.8	4600	23	ABLI22914	Drosophila melanog
718	16	0.8	2854	22	AAV18607	Human nervous syst	c 791	16	0.8	4601	21	AAV97542	Streptomyces albul
719	16	0.8	2885	22	AAV159008	Human polynucleoti	c 792	16	0.8	4641	23	ABLI16126	Drosophila melanog
720	16	0.8	2890	18	AAV89627	Genomic DNA encodi	c 793	16	0.8	4652	22	AAV160150	Human polynucleoti
721	16	0.8	2936	22	AAV21832	Human collagen gen	794	16	0.8	4705	23	ABLI10758	Drosophila melanog
722	16	0.8	2988	18	AAV84245	Rat ICAM-4 cDNA.	c 795	16	0.8	4718	23	ABLI21466	Drosophila melanog
723	16	0.8	2988	19	AAV34657	Rat intercellular	c 796	16	0.8	4771	23	ABLI22038	Drosophila melanog
724	16	0.8	2988	19	AAV11655	Rat ICAM-4 cDNA.	c 797	16	0.8	4783	23	ABLI09156	Drosophila melanog
725	16	0.8	2988	19	AAV19326	Rat ICAM-4 cDNA.	c 798	16	0.8	4800	19	AAV1658	Rat ICAM-4 genomic
726	16	0.8	2988	20	AAV36486	Rat ICAM-4 coding	c 799	16	0.8	4900	19	AAV19329	Rat ICAM overlappi
727	16	0.8	2988	20	AAV08991	DNA encoding novel	c 800	16	0.8	4900	20	AAV36488	Rat ICAM-1 coding
728	16	0.8	2997	23	AAV59359	S. epidermidis gen	c 801	16	0.8	4900	22	AAV36488	Partial rat ICAM-R
729	16	0.8	3016	22	AAV54928	Staphylococcus aur	c 802	16	0.8	4938	23	ABLI04108	Drosophila melanog
730	16	0.8	3025	18	AAV74489	Staphylococcus aur	c 803	16	0.8	5077	19	AAV1661	Rat cDNA clone for
731	16	0.8	3062	22	AAV54645	S. epidermidis gen	c 804	16	0.8	5077	19	AAV1661	Rat ICAM-4 clone E
732	16	0.8	3118	23	ABLI04190	Drosophila melanog	c 805	16	0.8	5077	19	AAV1661	Rat ICAM-4 partial
733	16	0.8	3121	22	AAV54245	S. epidermidis gen	c 806	16	0.8	5077	19	AAV19332	Rat ICAM coding se
734	16	0.8	3204	22	AAV94369	Human full-length	c 807	16	0.8	5077	20	AAV36491	Rat ICAM-4 coding
735	16	0.8	3223	23	AAV593540	DNA encoding novel	c 808	16	0.8	5163	24	ABLI33248	Human immune syste
736	16	0.8	3252	23	ABLI18460	Drosophila melanog	c 809	16	0.8	5176	19	AAV04017	Human multiple mye
737	16	0.8	3324	22	AAH16336	Human cDNA sequenc	810	16	0.8	5215	20	AAV91105	Group B Streptococ
738	16	0.8	3324	22	AAH16336	Human cDNA sequenc	811	16	0.8	5217	21	AAV56330	CKS-P66-CKS fusion
739	16	0.8	3347	21	AAV61121	SEN virus genome f	812	16	0.8				

C 813	16	0.8	5347	24	AAS94844	Human DNA sequence
C 814	16	0.8	5470	13	AAQ38596	Human glucose regu
C 815	16	0.8	5823	24	ABL12104	Drosophila melanog
C 816	16	0.8	5839	22	AAV46335	Hydra head activat
C 817	16	0.8	5976	22	AAS46301	Tumour suppressor
C 818	16	0.8	5983	23	AAS74036	DNA encoding novel
C 819	16	0.8	5987	24	ABL33563	Human immune syste
C 820	16	0.8	5997	22	AA166084	Leishmania major h
C 821	16	0.8	6016	22	AAH57390	Human skeletal mus
C 822	16	0.8	6096	22	AAS26766	Human genomic DNA
C 823	16	0.8	6171	19	AAV52170	Streptococcus pneu
C 824	16	0.8	6398	23	ABL25726	Drosophila melanog
C 825	16	0.8	6508	23	ABL09474	Drosophila melanog
C 826	16	0.8	6618	22	ABA18101	Human nervous syst
C 827	16	0.8	6666	12	AAQ10416	Mutant protease ge
C 828	16	0.8	6675	12	AAQ10414	Mutant protease ge
C 829	16	0.8	6675	12	AAQ10415	Mutant protease ge
C 830	16	0.8	6675	12	AAQ10411	Mutant protease ge
C 831	16	0.8	6675	12	AAQ10412	Mutant protease ge
C 832	16	0.8	6675	12	AAQ10413	Mutant protease ge
C 833	16	0.8	6675	12	AAQ10417	Mutant protease ge
C 834	16	0.8	6693	12	AAQ10871	Mutant protease ge
C 835	16	0.8	6711	12	AAQ10870	Mutant protease ge
C 836	16	0.8	6806	18	AAV74783	Staphylococcus aur
C 837	16	0.8	6882	22	ABA14861	Human nervous syst
C 838	16	0.8	6882	22	ABA15569	Human nervous syst
C 839	16	0.8	6979	23	AAS81429	DNA encoding novel
C 840	16	0.8	6990	22	AA158278	Human polynucleoti
C 841	16	0.8	7027	22	AA160064	Human LXR11 longer
C 842	16	0.8	7081	18	AAT74201	Staphylococcus aur
C 843	16	0.8	7104	23	AAS51998	Staphylococcus aur
C 844	16	0.8	7107	19	AAS4654	DNA encoding novel
C 845	16	0.8	7148	19	AAV32371	Staphylococcus mel
C 846	16	0.8	7148	24	ABA94251	Nucleotide sequenc
C 847	16	0.8	7156	24	ABA94251	Nucleotide sequenc
C 848	16	0.8	7156	20	AAK12966	Enterococcus faeca
C 849	16	0.8	7337	20	AAK13079	Enterococcus faeca
C 850	16	0.8	7389	22	AAS27788	DNA encoding novel
C 851	16	0.8	7417	23	ABL13024	Drosophila melanog
C 852	16	0.8	7434	23	AAS52179	Staphylococcus aur
C 853	16	0.8	7437	23	AAS55232	Staphylococcus aur
C 854	16	0.8	7449	24	ABL32276	Human immune syste
C 855	16	0.8	7469	19	AAV32372	Complete sequence
C 856	16	0.8	7469	21	AAAS9043	Nucleotide sequenc
C 857	16	0.8	7469	24	ABA94252	Nucleotide sequenc
C 858	16	0.8	7567	23	ABL12481	Drosophila melanog
C 859	16	0.8	7584	24	AAS95251	Long terminal repe
C 860	16	0.8	7726	22	AA166083	Leishmania donovan
C 861	16	0.8	7737	23	ABL02204	Drosophila melanog
C 862	16	0.8	7865	23	ABL06402	Drosophila melanog
C 863	16	0.8	7902	23	ABL06398	Drosophila melanog
C 864	16	0.8	7958	21	AAZ34835	Mouse integrin sub
C 865	16	0.8	7960	21	AAAS9072	Nucleotide sequenc
C 866	16	0.8	7960	24	ABA94274	Nucleotide sequenc
C 867	16	0.8	7989	21	AAAS9075	Nucleotide sequenc
C 868	16	0.8	7989	24	ABA94277	Nucleotide sequenc
C 869	16	0.8	8155	18	AAV74374	Staphylococcus aur
C 870	16	0.8	8365	20	AAZ20056	Plasmodium falcipa
C 871	16	0.8	8370	22	AAS46713	Tumour suppressor
C 872	16	0.8	8383	21	AAAS9071	Nucleotide sequenc
C 873	16	0.8	8383	24	ABA94273	Nucleotide sequenc
C 874	16	0.8	8471	22	AAS32667	Human genomic DNA
C 875	16	0.8	8480	20	AAZ27521	Interleukin-2 rece
C 876	16	0.8	8484	21	AAAS9091	Nucleotide sequenc
C 877	16	0.8	8748	23	ABL12592	Drosophila melanog
C 878	16	0.8	8821	23	ABL10984	Drosophila melanog
C 879	16	0.8	8907	22	ABA21312	Human nervous syst
C 880	16	0.8	8907	22	ABA21313	Human nervous syst
C 881	16	0.8	8907	22	AA103164	Human reproductive
C 882	16	0.8	8907	22	AA103165	Human reproductive
C 883	16	0.8	8930	19	AAV22834	Haemophilus paraga
C 884	16	0.8	8962	24	ABL32686	Human immune syste
C 885	16	0.8	9001	22	AAK84675	Human immune/haema
C 886	16	0.8	9120	22	ABA14655	Human nervous syst
C 887	16	0.8	9120	22	ABA14656	Human nervous syst
C 888	16	0.8	9180	22	AA104447	Human reproductive
C 889	16	0.8	9206	24	ABL33675	Human immune syste
C 890	16	0.8	9319	21	AAK20995	Human low adenosin
C 891	16	0.8	9319	21	AAK34873	Human adenosine re
C 892	16	0.8	9345	23	ABL09624	Drosophila melanog
C 893	16	0.8	9741	23	ABL33322	Human immune syste
C 894	16	0.8	9789	17	AA141852	cDNA encoding plas
C 895	16	0.8	10133	24	ABL32458	Human immune syste
C 896	16	0.8	10218	22	AAK77457	Human immune/haema
C 897	16	0.8	10296	23	ABL15024	Drosophila melanog
C 898	16	0.8	10610	19	AAV32375	Complete sequence
C 899	16	0.8	10610	21	AAAS9051	Nucleotide sequenc
C 900	16	0.8	10610	24	ABA94260	Nucleotide sequenc
C 901	16	0.8	10694	23	ABL07242	Drosophila melanog
C 902	16	0.8	10776	23	ABL03804	Drosophila melanog
C 903	16	0.8	10835	22	AAK79845	Human immune/haema
C 904	16	0.8	10982	22	AAS46264	DNA encoding novel
C 905	16	0.8	11034	24	AAV74407	Sheep alpha (1, 3)
C 906	16	0.8	11050	18	AAV74407	Staphylococcus aur
C 907	16	0.8	11176	22	AAQ1080	Mouse fosB DNA. M
C 908	16	0.8	11176	24	AAK22528	Mouse fosB DNA. M
C 909	16	0.8	11532	22	ABA07434	Human pancreatic c
C 910	16	0.8	11532	22	AAK91165	Human digestive sy
C 911	16	0.8	12103	22	AA105045	Human reproductive
C 912	16	0.8	12103	22	AAS33419	DNA encoding human
C 913	16	0.8	12225	18	AAK79257	Human by gene (sho
C 914	16	0.8	12278	22	ABA14585	Human nervous syst
C 915	16	0.8	12278	22	ABA14664	Human nervous syst
C 916	16	0.8	12283	12	AAQ10353	Nucleotide sequenc
C 917	16	0.8	12343	22	AAS46240	DNA encoding novel
C 918	16	0.8	12460	24	ABL32069	Human immune syste
C 919	16	0.8	12616	18	AAK69256	Human by gene (lon
C 920	16	0.8	12789	24	AAK16876	Human interleukin
C 921	16	0.8	13003	23	ABL12480	Drosophila melanog
C 922	16	0.8	13605	22	AAK37081	Human musculoskele
C 923	16	0.8	13767	23	ABL15058	Drosophila melanog
C 924	16	0.8	14333	22	AAK79846	Human immune/haema
C 925	16	0.8	14455	19	AAV32374	Complete sequence
C 926	16	0.8	14455	21	AAAS9050	Nucleotide sequenc
C 927	16	0.8	14455	24	ABA94259	Nucleotide sequenc
C 928	16	0.8	14458	23	ABL10220	Drosophila melanog
C 929	16	0.8	14537	22	AAS46355	Tumour suppressor
C 930	16	0.8	14657	23	ABL18606	Drosophila melanog
C 931	16	0.8	14775	24	AB199535	Mouse ischaemic co
C 932	16	0.8	15644	22	AAS34611	Human DNA for a no
C 933	16	0.8	16870	20	AAK13035	Enterococcus faeca
C 934	16	0.8	16891	20	AAK37084	MEFV gene sequence
C 935	16	0.8	17674	24	ABL33344	Human immune syste
C 936	16	0.8	18177	10	AAAG0490	DNA of human retin
C 937	16	0.8	18303	20	AAK04502	Human retinoblasto
C 938	16	0.8	18475	19	AAV52171	Streptococcus pneu
C 939	16	0.8	18598	24	ABL32386	Human immune syste
C 940	16	0.8	18869	23	ABL04818	Drosophila melanog
C 941	16	0.8	20752	22	AAK75098	Human immune/haema
C 942	16	0.8	21591	20	AAK13047	Enterococcus faeca
C 943	16	0.8	21732	22	AAS39820	Genomic sequence #
C 944	16	0.8	21732	22	AAK30176	Human digestive sy
C 945	16	0.8	21799	22	ABA07343	Human pancreatic c
C 946	16	0.8	21799	22	AAS32770	Human immune/haema
C 947	16	0.8	22401	22	AAK71911	Human immune/haema
C 948	16	0.8	22402	22	AAK71912	Human immune/haema
C 949	16	0.8	23054	22	AAK84676	Human immune/haema
C 950	16	0.8	23452	22	AAS42122	Genomic sequence #
C 951	16	0.8	23457	22	AAS42121	Genomic sequence #
C 952	16	0.8	23458	22	AAS42120	Genomic sequence #
C 953	16	0.8	24259	22	AAS46691	Tumour suppressor
C 954	16	0.8	24699	22	AA136054	Human musculoskele
C 955	16	0.8	25003	22	ABA19679	Human nervous syst
C 956	16	0.8	25131	22	AAK84544	Human immune/haema
C 957	16	0.8	26370	23	ABL07082	Drosophila melanog
C 958	16	0.8	26879	23	AAS59524	Propionibacterium

959	16	0.8	27433	23	ABL20312	Drosophila melanog
960	16	0.8	28866	20	AAV22304	Human IL-1ra BAC c
c 961	16	0.8	28918	23	ABL14456	Drosophila melanog
c 962	16	0.8	31183	22	AAD03963	Adenovirus Ag-5 de
c 963	16	0.8	31446	21	AAA09088	AdB-beta galactos
964	16	0.8	31741	22	ABA21231	Human nervous syst
c 965	16	0.8	32026	18	AAI60559	Recombinant adenov
c 966	16	0.8	32165	21	AAA09092	AdMMTV-beta-galact
c 967	16	0.8	32165	21	AAA14723	Nucleotide sequenc
c 968	16	0.8	32166	21	AAA09090	AdPSA-beta-galact
c 969	16	0.8	32166	22	AAAC89170	AdSVphxHDE region
c 970	16	0.8	32167	21	AAA14803	Nucleotide sequenc
c 971	16	0.8	32167	21	AAZ93332	Partial sequence o
c 972	16	0.8	32172	22	AAI03473	Human reproductive
c 973	16	0.8	32184	22	AAI05850	Human reproductive
974	16	0.8	32196	22	ABA18857	Human nervous syst
c 975	16	0.8	32204	22	AAI05849	Human reproductive
976	16	0.8	32220	22	ABA20618	Human nervous syst
c 977	16	0.8	32886	21	AAA09086	AdRSV-beta-galacto
c 978	16	0.8	33592	22	AAAC85018	Adenovirus anti-ca
979	16	0.8	33655	23	ABL08422	Drosophila melanog
c 980	16	0.8	33659	22	AAAC85020	Adenovirus anti-ca
c 981	16	0.8	33988	22	AAAC85023	Adenovirus anti-ca
c 982	16	0.8	34302	21	AAZ94163	Adenovirus vector
c 983	16	0.8	34303	19	AAV07261	Adenoviral vector
c 984	16	0.8	34341	22	AAAC85019	Adenovirus anti-ca
c 985	16	0.8	34382	20	AAI15627	Recombinant adenov
c 986	16	0.8	34427	20	AAV07371	Adenovirus vector
c 987	16	0.8	34448	22	AAAC85021	Adenovirus anti-ca
c 988	16	0.8	34658	22	AAK69489	Human immune/haema
c 989	16	0.8	34737	22	AAAC85024	Adenovirus anti-ca
c 990	16	0.8	35000	18	AAI60557	Recombinant cis-ac
c 991	16	0.8	35408	18	AAI59272	Adenovirus anti-ca
c 992	16	0.8	35724	19	AAAC85022	Adenovirus anti-ca
c 993	16	0.8	35935	19	AAV07258	Adenovirus 5 genom
c 994	16	0.8	35935	22	AAAC85026	Complete nucleotid
c 995	16	0.8	36114	22	AAAC85025	Adenovirus anti-ca
c 996	16	0.8	36538	18	AAI60558	Recombinant trans-
c 997	16	0.8	37339	22	AAI15612	Mouse osteocalcin
998	16	0.8	38682	22	AAI21770	Human gene for col
c 999	16	0.8	40349	21	AAI22278	BAC containing rep
1000	16	0.8	43095	21	AAI68254	Bacteriophage 3A c

ALIGNMENTS

RESULT 1

AAV27323 standard; DNA; 1999 BP.

AAV27323;

02-OCT-1998 (first entry)

Streptococcus pneumoniae SP001 nucleotide.

Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;

detection; pneumonia; otitis media; meningitis; ss.

Streptococcus pneumoniae.

Location/Qualifiers

2..1999

/tag= a

/product= "SP001"

/note= "no stop codon given"

30-OCT-1997; 97WO-US19422.

XX	31-OCT-1996;	96US-0029960.
PR	(HUMA-) HUMAN GENOME SCI INC.	
XX	Choi GH, Hromocky J A, Johnson LS, Kunsch CA;	
PA	WPI; 1998-272224/24.	
XX	P-PSDB; AAW5063.	
DR		
XX		
PT	Nucleic acid encoding antigenic peptide(s) from Streptococcus	
PT	pneumoniae - or their epitope-containing fragments, useful in	
XX	protective or therapeutic vaccines, and for diagnosis	
PS	Claim 1; Page 48; 118pp; English.	
XX		
CC	The present sequence encodes a protein from Streptococcus pneumoniae.	
CC	The nucleic acid sequence encoding the Streptococcus pneumoniae protein	
CC	can be useful in vaccines for inducing protective antibodies against	
CC	Streptococcus pneumoniae, for treatment or prevention of infection e.g.	
CC	pneumonia, otitis media or meningitis. Probes based on the nucleic acid	
CC	are used to detect Streptococcus infection (by usual hybridisation or	
CC	amplification methods), also for isolating Streptococcus genes or their	
CC	allelic variants. The protein can be used similarly for detecting specific	
CC	antibodies in standard immunoassays, especially for diagnosing or	
CC	monitoring infections. Antibodies which bind the protein are used to	
CC	detect corresponding antigens, to purify the protein and for passive	
CC	immunisation (optionally coupled to a toxin). Vaccines are administered,	
CC	e.g. by injection, orally or through the skin, typically at 0.01-1000	
CC	(especially 10-300) mu g/ml per dose.	
XX		
SQ	Sequence 1999 BP; 656 A; 485 C; 376 G; 482 T; 0 other;	

Query Match 100.0%; Score 1999; DB 19; Length 1999;

Best Local Similarity 100.0%; Pred No. 0;

Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TAAATCTACGACATATAATCACTCATCTGCTGACTTGGTTGACGCGCGTCAA	60
DB	1	taaatctacgacataataataatcaactcatctgcttgcgtcgaacgcgctcaaa	60
QY	61	TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATGTTTCTATCGAAGCA	120
DB	61	tgcccaagctaagataltccacagatttgtaaggaatcglttctatcgaagacca	120
QY	121	TGCGTCTTGACACAGGAGGATGATGATACATCCGATCTGGAGCTTCTTGCGCAA	180
DB	121	tcgctctcttgaccacaggggagattgataccaccgatactccgagagcttcttcgcaa	180
QY	181	TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTACCCCAAGTTGATTAAGTTGAC	240
DB	181	tctgcaaaagaatctccctcccaaggatgataacttccaccacagttgataagttgac	240
QY	241	TTTACTTTTCACTTCGACTTCGACCAAGACTATTTCTCGTAAGGCTCAGAACCTTGTT	300
DB	241	ttactttcaacttcgacttcgacacagactatttctcgaagctcgaagaacttggtt	300
QY	301	AGGATTCAGTTGACCAAAAGCAACCAAGCAAGAAAGTTCACCTACTATTAATAA	360
DB	301	aggatattcagttagaacaaaagaacacgaagaagaatttgaacctactataataaa	360
QY	361	GGTCTACATGCTTAATGGGAACATGGAATGCAAGACAGCTCAAAACATCATGTTAA	420
DB	361	ggtctacatgcttaaatgggaactatggaatgacagacgctcaaaactactatgtaa	420
QY	421	AGACGTCATATTAATTAAGTTTACCTCAGTTAGCTTGCTGGAATGCTCAGGCAAC	480
DB	421	agacgtcaatatttaagttaagttacctcagttagcttgctgagatgctcagacacc	480
QY	481	AAACCAATATGACCCCTATTCACATCCAGAGACGCCCAAGAGCGCGCAACTTGCTT	540
DB	481	aaaccaatatgacccctattcacatccagaagagcaccgaacgcgcgaacttgltct	540

OY 541 ATCTGAATGAAAAATCAAGGCTACATCTGCTGAACAGTATGAGAAAGCACTCAATAC 600
 |||||
 Db 541 atctgaatgaaaaaatcaaggtacatctctgctgaacagtatgagaagaagcagtcataac 600
 OY 601 ACCAATTACTGATGACTCAAAAGTCTCAAAATCAGCAAGTAATTACCTGCTTACATGGA 660
 |||||
 Db 601 accaattactgatgactcaaaagtctcaaatcaagtaattaccctgcttaccatgga 660
 OY 661 TTAATTCCCAAGSAGTATCATCAATCAAGTATGAGAAAGAAAGGCTATATACCTACTCAC 720
 |||||
 Db 661 ttaattcccaagsagtatcatcaatcaagtatgagaagaagaacagctataaactactcac 720
 OY 721 AACTGGGATGATGCTTACACAATGTAGAACCAAGAGCTCAAAAACATCTGGGATAT 780
 |||||
 Db 721 aactgggatgcttaccacaatgtagaaaccaagagctcaaaaacatctgggatat 780
 OY 781 TTACAAATACAGACGAATACGTTGCTTATCCAGACGATGAATTCGAAGTCGCTTACCAT 840
 |||||
 Db 781 ttacaaatacagacgaatacgttgcttattccagacgattgaatttcgaagtcgcttaccat 840
 OY 841 TGTGTGATGTTTGAAGTCAATGCGCAGTACGAGACGCGCATACAGTCAAGTAA 900
 |||||
 Db 841 tgtgtgatgtttgaagtcaatgctgcttgaagacgagcgcgcatacagtcgaagtaa 900
 OY 901 TGTTCCTTCGGAATTAACCAAGCAGTACGAACAAACCGCAGTGGGATCAACTATGAA 960
 |||||
 Db 901 tgttccttcggaatttaaccaagcagtagaacaacacgagctgggatcaactatgaa 960
 OY 961 ACCGATACAGACTATGCTCTGCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT 1020
 |||||
 Db 961 accgatcacagactatgctctgcttggagtacgggtgtctacgattccaactgctactat 1020
 OY 1021 CGTTCCGATGAGCCCTATACCTACCCGTGGACAATACCTCCGTTTATAACGGGATAG 1080
 |||||
 Db 1021 cgttccgatgagccctatacctacccgtggacaatactccgttataaacgggatatag 1080
 OY 1081 GGGCTACTTTGGCAACATCACCCTTGCATATACGCCCTGCAACAATCGCAAAACGTCACAC 1140
 |||||
 Db 1081 gggctactttggcaaatcacccttgcatactacgcccgtgacaacaatcgcaaaacgtccacag 1140
 OY 1141 CGTGAACACTTAACAGAGTGGGACTCAACCGCGCAAGACTTTCCTTAATGCTTAGG 1200
 |||||
 Db 1141 cgtgaacacttaaacagaagtgctgactcaaccgcgcaagacttcccttaattgcttagg 1200
 OY 1201 AATCGACTCCCAAGTATTCACACTCAATGGCAATTCGAATACACAACGGAATCAGA 1260
 |||||
 Db 1201 aatcgactcccaagtattcacactcaaatgccaattcaagaacacacacggaatcaga 1260
 OY 1261 CAAAAAATATGAGCAAGTAGTGAAGAAAGATGGCTGCTTACGCTGCTTGCANAATG 1320
 |||||
 Db 1261 caaaaaatatgagcaagtagtgaagaagtgctgcttgaacgcttgccttgcgaatag 1320
 OY 1321 TGGACTTACTTAACCAATGTATTCATTAAGTGTCTTATAGTGAAGGAGTGAAGA 1380
 |||||
 Db 1321 tggacttactttaaccaatgtatcatataaagtcgtctttagtgaaggagtgaaaga 1380
 OY 1381 AGAGTCTCTTAATGTGCAACTCGTGCATGAAGAAAGACAGCTATATGATGACCGA 1440
 |||||
 Db 1381 agaagttcttaatgtcgcaactcgycatgagaagaacagcagctataatgctgcga 1440
 OY 1441 CATGATGAAAACAGTCTTACCTTATGGAAGTGAAGAAATGCCCTATCTGCTGGCTCC 1500
 |||||
 Db 1441 catgatgaaaacagcttctgactatggaactggaagaatgctactctgctgctcc 1500
 OY 1501 TCAGGCTGGTAAACAGAACTCTTACTATACAGACGAGGAATTTGAAAACACATCA 1560
 |||||
 Db 1501 tcaggctggtaaacagaaactcttaactatacagacggaattgaaacacacatcaa 1560
 OY 1561 GACCTTCATTTTGTAGACGCTGATGACTATTTGCTGGCTTATAGCGGTAATATTCAT 1620
 |||||
 Db 1561 gaccttcatttgtagacacctgagaaactattgctgctatacgcgtaaatatltcaat 1620

OY 1621 GGCTGATGAGACAGGCTATTTCAACCGTCTGACACCACTTTGAGGAATGGCCTTACGGT 1680
 |||||
 Db 1621 ggctgatgagacaggtatttcaacgctctgacacaccttggaggaatggccttacgg 1680
 OY 1681 CGCTGCCAAAGTTTACCCTCTATGATGACTTACCTGCTGGAAGGAAGCAATCAGAA 1740
 |||||
 Db 1681 cgctgccaaagtttaccctctatgacttaccctgctggaaggaaagcaatcagaaga 1740
 OY 1741 TTGGAATATACAGAGGGGCTCAGCAAAATGAGAAATTCGTTTAAAAATGTCGCTCG 1800
 |||||
 Db 1741 ttggaatatacagagggtctcacagaatggaatctgatatctaaaaatggctcg 1800
 OY 1801 TTCTACGTGGAACCTCACTGCTCTCAGACAAACCCCATCAACTCAAGTCAAGCTATC 1860
 |||||
 Db 1801 ttctacgtggaacctcactgctccacaacaaccccatcaactcaagttcaagctatc 1860
 OY 1861 ATCAGATAGTTCACTTACAGCTTCTAGCTTACCTCAGCAACAAATATATAGTACGAC 1920
 |||||
 Db 1861 atcagatagttcaacttcaagcttagctcaaccactccaagcacataatagtcagac 1920
 OY 1921 TACCAATCCTAACAAATATACGACAAATACCAATACACCCCTGATCAACAAATCGAA 1980
 |||||
 Db 1921 taccaatccctaacaataatacgcacaatacaatacaacccctgatacaaaaatcaga 1980
 OY 1981 TCCTCAACGACGACAACCA 1999
 |||||
 Db 1981 tcctcaacgacgacaacca 1999

RESULT 2
 AAV52278/c
 ID AAV52278 standard; DNA; 10711 BP.

AAV52278;

23-OCT-1998 (first entry)

Streptococcus pneumoniae genome fragment SEQ ID NO:145.

Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay; computer readable medium; vaccine; pharmaceutical composition; ds.

Streptococcus pneumoniae.

MO9818931-A2.

07-MAY-1998.

30-OCT-1997; 97WO-US19588.

31-OCT-1996; 96US-0029960.

(HUMA-) HUMAN GENOME SCI INC.

Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M; Kunsch CA, Rosen CA;

WPI; 1998-272225/24.

Computer-readable medium with recorded Streptococcus pneumoniae polynucleotide sequences - useful in diagnostic kits and assays, and pharmaceutical compositions and vaccines for Streptococcus pneumoniae

Claim 1; Page 972-978; 1409pp; English.

The present invention describes a computer readable medium which has the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524) recorded on it, or a representative fragment or a sequence at least 95% identical to SEQ ID NO:1 to 391. The nucleotide sequences depicted in SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from Streptococcus pneumoniae. The present invention also describes an isolated nucleic acid molecule encoding a homologue of any of the

CC fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the
CC nucleic acid library is produced by a process comprising: (a) screening
CC a genomic DNA library using as a probe a target sequence defined by any
CC of the sequences in SEQ ID NO:1 to 391, identifying members of the
CC library which contain sequences that hybridise to the target sequence and
CC isolating the nucleic acid molecules from the members; or (b) isolating
CC mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
CC molecules whose nucleotide sequence is homologous to amplification
CC primers derived from the fragment of the S. pneumoniae genome to prime
CC the amplification and isolating the amplified sequences; The computer
CC readable medium can be used in a computer-based system for identifying
CC fragments of the S. pneumoniae genome of commercial importance, or
CC expression modulating fragments of the S. pneumoniae genome. Products
CC from the present invention can be used in diagnosis kits and assays, and
CC pharmaceutical compositions and vaccines for S. pneumoniae.

XX Sequence 10711 BP; 3280 A; 2056 C; 2412 G; 2963 T; 0 other:

Query Match 100.0%; Score 1999; DB 19; Length 10711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAAATTAATACTCACTCTGACTGGGTTCTGACGCCGCTCAA 60
DB 9767 TAAATCTACGACAAATTAATACTCACTCTGACTGGGTTCTGACGCCGCTCAA 9708
QY 61 TGCCCAAGTAATGATTTCCACAGATTTGGTTAAGCAATCGTTCTATCGAAGCA 120
DB 9707 TGCCCAAGTAATGATTTCCACAGATTTGGTTAAGCAATCGTTCTATCGAAGCA 9648
QY 121 TCGCTTCTTGACACACAGGGGATTTGATACATCCGTATCCTGGAGCTTCTTGCCAA 180
DB 9647 TCGCTTCTTGACACACAGGGGATTTGATACATCCGTATCCTGGAGCTTCTTGCCAA 9588
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTCACCAACAGTTGATTAAGTTGAC 240
DB 9587 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTCACCAACAGTTGATTAAGTTGAC 9528
QY 241 TTACTTTTCACTTCGACTCCGACGACATTTCTCGTAAGGCTGAGAGGTTGGTT 300
DB 9527 TTACTTTTCACTTCGACTCCGACGACATTTCTCGTAAGGCTGAGAGGTTGGTT 9468
QY 301 AGCGATTCACTTAGAACAAAAGAACCAAGCAAGAAATTTGACTATATATAATA 360
DB 9467 AGCGATTCACTTAGAACAAAAGAACCAAGCAAGAAATTTGACTATATATAATA 9408
QY 361 GGTCTACATCTTAATGGGAACCTATGGAATGACAGACAGCTCAAACTATATGTTAA 420
DB 9407 GGTCTACATCTTAATGGGAACCTATGGAATGACAGACAGCTCAAACTATATGTTAA 9348
QY 421 AGACCTCAATTAATTTAGTTTACCTAGCTAGCTTCTGCTGGAGTGGCTCGAGCAC 480
DB 9347 AGACCTCAATTAATTTAGTTTACCTAGCTAGCTTCTGCTGGAGTGGCTCGAGCAC 9288
QY 481 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGCCGAAACTTGCTTT 540
DB 9287 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGCCGAAACTTGCTTT 9228
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTGCTGAACAGTATGAGAAGCAGTCAATAC 600
DB 9227 ATCTGAAATGAAAAATCAAGGCTACATCTGCTGAACAGTATGAGAAGCAGTCAATAC 9168
QY 601 ACCAATTAATGAGACTACAAAGTCAATACAGCAAGTAATTAACCTGCTTATCATGSA 660
DB 9167 ACCAATTAATGAGACTACAAAGTCAATACAGCAAGTAATTAACCTGCTTATCATGSA 9108
QY 661 TAATTAACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGAACAGGCTATACCTACTCAC 720
DB 9107 TAATTAACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGAACAGGCTATACCTACTCAC 9048
QY 721 AACTGGAGTGAATGTCTACCAAAATGTAGACCAAGAGCTCAAAACATCTGTGGATAT 780
DB 721 AACTGGAGTGAATGTCTACCAAAATGTAGACCAAGAGCTCAAAACATCTGTGGATAT 780

DB 9047 AACTGGAGTGAATGTCTACCAAAATGTAGACCAAGAGCTCAAAACATCTGTGGATAT 8988
QY 761 TTACAATACAGCAATACGTTGGCTTATCCAGACGATGAATGCAAGCTGCTTACAT 840
DB 8987 TTACAATACAGCAATACGTTGGCTTATCCAGACGATGAATGCAAGCTGCTTACAT 8928
QY 841 TGTGTAGTGTCTTAACGTAAGTAAGTCAATTTGCCACCTTGAGAGCACGCCATGATCAAGTAA 900
DB 8927 TGTGTAGTGTCTTAACGTAAGTAAGTCAATTTGCCACCTTGAGAGCACGCCATGATCAAGTAA 8868
QY 901 TGTTCCTTTGGGAATTAACCAAGCAGTGAAGAACAAACCCGAGCTGGGATCACTATGAA 960
DB 8867 TGTTCCTTTGGGAATTAACCAAGCAGTGAAGAACAAACCCGAGCTGGGATCACTATGAA 8808
QY 961 ACCGATCACAGATATAGCTCTGCTGGAGTACGAGTGTATGATTAACCTGCTACTAT 1020
DB 8807 ACCGATCACAGATATAGCTCTGCTGGAGTACGAGTGTATGATTAACCTGCTACTAT 8748
QY 1021 CGTTACAGATGAGCCCTATTAACCTACCTGGGACAAATACTCTGTTATTAACCTGGATAG 1080
DB 8747 CGTTACAGATGAGCCCTATTAACCTACCTGGGACAAATACTCTGTTATTAACCTGGATAG 8668
QY 1081 GGGCTACTTTGGCAACATCACTCTGCAATACGCCCTGCAACATCGGGAACGCTCCAGC 1140
DB 8687 GGGCTACTTTGGCAACATCACTCTGCAATACGCCCTGCAACATCGGGAACGCTCCAGC 8628
QY 1141 CGTGAACACTTAACCAAGGTGGAGCTCAACCGGCAAGACTTCTCTTAATGGTCTAGG 1200
DB 8627 CGTGAACACTTAACCAAGGTGGAGCTCAACCGGCAAGACTTCTCTTAATGGTCTAGG 8568
QY 1201 AATGACTACCCAGTATTCAGTCACTCAATTCGCAATTTCAAGTAACACCAACCGAATGAG 1260
DB 8567 AATGACTACCCAGTATTCAGTCACTCAATTCGCAATTTCAAGTAACACCAACCGAATGAG 8508
QY 1261 CAAAAAATATGAGCAAGTATGTAAGATGCGCTGCTTACGCTCTTGGAAATGG 1320
DB 8507 CAAAAAATATGAGCAAGTATGTAAGATGCGCTGCTTACGCTCTTGGAAATGG 8448
QY 1321 TGAAGCTTACTATTAACCAATGTATTCATCAATTAAGTGTCTTATGATGAGTGA 1380
DB 8447 TGAAGCTTACTATTAACCAATGTATTCATCAATTAAGTGTCTTATGATGAGTGA 8368
QY 1381 AGAGTTCTTAATGTGGAAGCTGTGCGCATGAAGAAAGACAGCCTATATGATGACGA 1440
DB 8387 AGAGTTCTTAATGTGGAAGCTGTGCGCATGAAGAAAGACAGCCTATATGATGACGA 8328
QY 1441 CATGATAAAAACAAGTCTTGAATTTGGAATCTGGAACGAAATGCCATATTTGCTGGCTCCC 1500
DB 8327 CATGATAAAAACAAGTCTTGAATTTGGAATCTGGAACGAAATGCCATATTTGCTGGCTCCC 8268
QY 1501 TCAGGCTGTGTAACAGGAACCTCTAATCTATACAGAGAGGAAATTTGAAACACCATCAA 1560
DB 8267 TCAGGCTGTGTAACAGGAACCTCTAATCTATACAGAGAGGAAATTTGAAACACCATCAA 8208
QY 1561 GACCTCTCAATTTGTAGCAGCTGTATGATTAATTTGCTGCTATACGCGTAATATTTCAAT 1620
DB 8207 GACCTCTCAATTTGTAGCAGCTGTATGATTAATTTGCTGCTATACGCGTAATATTTCAAT 8148
QY 1621 GCGTGTATGAGACAGGCTATTTCAACCGCTGACACCACTTGTAGGCAATGGCTTACAGGT 1680
DB 8147 GCGTGTATGAGACAGGCTATTTCAACCGCTGACACCACTTGTAGGCAATGGCTTACAGGT 8088
QY 1681 CGCTGCCAAGATTACGCTCTATGATGAGACTACTCTGTGAAGGAGCAATCCAGAGA 1740
DB 8087 CGCTGCCAAGATTACGCTCTATGATGAGACTACTCTGTGAAGGAGCAATCCAGAGA 8028
QY 1741 TTGGAATATACAGAGGGGCTTACAGAAATGAGAAATTCGATTTAAAAATGGTCTCG 1800
DB 8027 TTGGAATATACAGAGGGGCTTACAGAAATGAGAAATTCGATTTAAAAATGGTCTCG 7968
QY 1801 TTCTAGTGAAGTCACTGCTGCTCCACAAACCCCATCACTGAAAGTTCAAGCTCATC 1860
DB 7967 TTCTAGTGAAGTCACTGCTGCTCCACAAACCCCATCACTGAAAGTTCAAGCTCATC 7908

QY 1861 ATGAGTAGTTCACCTTACAGTCTAGTCAACCTCCAGACAAATATAGTACGAC 1920
|||||
DB 7907 ATGAGTAGTTCACCTTACAGTCTAGTCAACCTCCAGACAAATATAGTACGAC 7848
QY 1921 TACCAATCTTAACAAATATACGACAAATCAATATACACCCCTGATCAACAAATACGAA 1980
|||||
DB 7847 TACCAATCTTAACAAATATACGACAAATCAATATACACCCCTGATCAACAAATACGAA 7788
QY 1981 TCCTCAACGACGACACCA 1999
|||||
DB 7787 TCCTCAACGACGACACCA 7769
RESULT 3
AAT08027
ID AAT08027 standard; DNA; 2048 BP.
AC AAT08027;
XX
XX 03-DEC-1996 (first entry)
DE S. pneumoniae penicillin binding protein 1A soluble variant DNA.
XX
XX Penicillin binding protein; PBP 1A; bifunctional protein;
KM transglycosylase; transpeptidase; identification; assay; inhibitor;
KM antibiotic resistant; bacteria; soluble variant; protein structure;
XX X-ray crystallography; determination; ds.
OS Streptococcus pneumoniae.
FH
FH Key Location/Qualifiers
FT CDS 1..2046
FT /tag= a
XX
XX GB2290792-A.
XX
XX 10-JAN-1996.
XX
XX 29-JUN-1995; 95GB-0013306.
XX
XX 24-NOV-1994; 94SE-0004072.
XX 01-JUL-1994; 94IN-0000580.
XX
XX (ASTR) ASTRA AB.
XX
XX Balganesch TS, Town CM;
PI
XX
XX WPI: 1996-042232/05.
DR P-PSDB; AAM04359.
XX
XX Sol. deriv. of bifunctional penicillin binding protein (BPBP) -
PT opt. lack transglycosylase activity, useful to identify and assay
PT for antibodies or cpds. which bind BPBP
XX
XX Claim 11; Pages 65-68; 108bp; English.
XX
XX The present sequence encodes a soluble variant of the S. pneumoniae
CC penicillin binding protein (PBP) 1A, which lacks the 1st 38 amino
CC acid residues of the wild type protein. Wild type PBP is a
CC bifunctional protein, which binds the cell membrane when expressed
CC in a bacterial cell, having transglycosylase and transpeptidase
CC activities. The variant protein (NCIMB 40665) in conjunction with
CC a labelled anti-bifunctional PBP monoclonal antibody, can be used
CC to identify and assay for cpds. which bind bifunctional PBP. Such
CC cpds., as inhibitors of bifunctional PBP have a potential use in
CC therapeutic cpds. which inhibit the growth of antibiotic resistant
CC bacteria. The soluble variant may also be used in X-ray
CC crystallography.
SQ Sequence 2048 BP; 667 A; 498 C; 390 G; 493 T; 0 other;

Query Match 71.9%; Score 1438; DB 17; Length 2048;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 TAAATCTACGACAAATTAATCACTATGCTGACTGGTTGTAAGCCGCTCAA 60
|||||
DB 48 taaactcagcaataaataactcattgcttgacttggctcgaacgcgcgtcaa 107
QY 61 TGGCCAGGTAATGATATCCGACAGATTGGTTAAGGCAATGGTTTATCGAAGACCA 120
|||||
DB 108 tgcacaagcctaaatgataatcccaagatttgtaaggaatcgtttctacgaagacca 167
QY 121 TCGCTTCTTCGACACAGGGGAGTTGATACATCCGTATCCCTGGAGCTTTCTTGGCAA 180
|||||
DB 168 tgcctcttcgaccacaggggagttgataccatcgcgtatcccttgagcgtttcttcgcaa 227
QY 181 TCTGCAAGCAATTCCTCCAGAGTGATCACTGCACCCAAAGTTGATTAAATTGAC 240
|||||
DB 228 tctgcaagaacatctccctcacaagtgatcagctcctcactcaacagtgatlaagttgac 287
QY 241 TTACTTTCACTTGCAGCTTCGACACAGACTATTTCTCGTAAGGCTCAGGACCTGGTT 300
|||||
DB 288 ttaactttcaactcgtgacttcgacagactatcttcglaaggtccaagaaagcttgglt 347
QY 301 AGCGATTCAAGTTAGAACAAACCAACCAAGCAAGAAATCTTACCTACTATTAATAA 360
|||||
DB 348 agcgattcagttagaacaaacaaagcaacgaagaatcttgactactataataaa 407
QY 361 GGTCTACATGCTTAATGGCACTATGGAATGCAGACAGCAAGCTCAAAACTATATGTTAA 420
|||||
DB 408 ggtctacatgcttaaatggaactatgaaatgcagacagctcaaaactactatgttaa 467
QY 421 AGACCTCAATTAATTAAGTTTACCTAGTTAGGCTTGGGGCTGGAATGCTCAGGACAC 480
|||||
DB 468 agaccccaaatlaactaagttaccctcagtaggcttgcgctggaatgcttcaagcacc 527
QY 481 AAACCAATATGACCCCTATTTCACATCCAGAGGACGCCAAGAGCGCGAAACTTGGTCTT 540
|||||
DB 528 aaaccaataatgaccctattcacaatccagaagcagcccaagcgcgaaacttggctt 587
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGTAACAGATATGAGAAAGCACTCAATAC 600
|||||
DB 588 atctgaatgaaaaatcaaggtcacatctctgctgaacgtatgaaagcaactacatc 647
QY 601 ACCAATTACTGATGACTCAAAAGTCTCAAAATAGCAAGTAATTAACCTCTTACATGGA 660
|||||
DB 648 accaattactgattggtctacaagttccaaatccaaagtaaatcattcgtctacatgga 707
QY 661 TAATTACTCAAGGAAGTCAATCAAGTTGAAGAAGAAACAGGCTATTAACCTACTACAC 720
|||||
DB 708 taattactcaaggaagtcataatcaatgaagtgagaagaaacagctataactactcacc 767
QY 721 AACTGGATGATGTTTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
|||||
DB 768 aactggatgattggttaccacaatgtagaccaagaagtcacaaaacatcttgggatat 827
QY 781 TTACATACAGACGATFACGTTGCTATCCAGAGATGAAATTCAGAGTGGCTCTACCAT 840
|||||
DB 828 ttacaatacagaagaatactgtgcttaccagaagcagatgtgcaagtcgcttaccacat 887
QY 841 TGTGATGTTTCAAGGTAAGTCAATGCTCCAGCTAGAGAGCAGCCATCAGTCAAGTAA 900
|||||
DB 888 tgtgattgttctcaaggttaaggtcattgcccagtcagagacgcacatcagtcagtaa 947
QY 901 TGTTCCTTCGGAATTAACCAAGCAAGTATGAAACAAACCGAGATGGGATCAACTATGAA 960
|||||
DB 948 tgttcttcctcgaaatcaacaaagcagtagaacaacacgcgacttgggatacactatgaa 1007
QY 961 ACCGATCAGAGACTATGCTGCTGCTGGAGTACGCTGATAGTTTCAACTCTCTAT 1020
|||||
DB 1008 accgatcaagactatgctcgccttggagtagcagtgctcagatcaactgcacat 1067
QY 1021 CGTTACAGATGAGCCCTATTAACCTACCTGGACAAATACTCCTGTTTATTAACGTGGATAG 1080

|||||
 Db 1068 cytcacagatgagccataactaccctgggacaataaccctgttataactgagtag 1127
 OY 1081 GGGCTACTTTGGCAACATCATCTGCAATACGGCCCTGCAACATCGGCCAGC 1140
 Db 1128 gggctactcttggcaacatcactcgtcgaatacgcctcgaacacgcgaaacgccccagc 1187
 OY 1141 CGTGAACACTCTAATAACAGGTGGGACTCAACCGGCCAGACGCTTCTTAATGGTCTAGG 1200
 Db 1188 cgtggaaccttaacaagaagtcgagctcaacgcgcgaagacttcttaattgtctcgg 1247
 OY 1201 AATGACTACCCAACTTCTACTACTCAATGCCATTTCAGTACACACCGAATCAGA 1260
 Db 1248 aatgactaccgaagtattactactcaaatgaccttcaagtaacaacacgaatcaga 1307
 OY 1261 CAAAAAATATGAGCAGTGTGTGAAGATGGGCTGCTAGCGCTGCTTGTGCAATGG 1320
 Db 1308 caaaaataatggaagaagtgagaaagatgagctgctgaagctgaccttgcgaatgg 1367
 OY 1321 TGAACCTTACTATTAACCAATGTATATCCATTAAGTGTCTTATGATGAGTGAATA 1380
 Db 1368 tggagactactataaacaatgatatccataaagtcgtctttagtgatggagtagaaa 1427
 OY 1381 AGAGTTCTTAATGTGGAACTGCTGCCATGAGAGAAAGCAGACGCTATATGATGACCGA 1440
 Db 1428 agagttctctaattgctggaactcgtccatgaagaagacagacgacctatgatgacgga 1487
 OY 1441 CATGATAAAAACATCTTGTGAGTGTGAACTGAGCAAAATCCCTATCTTGGTGGCTCCC 1500
 Db 1488 catgatgaaaaacagctcttgatgataagacggaagaatgacctatctgcttgagctccc 1547
 OY 1501 TCAGGCTGTGTAACAGGAACCTCTAATCTATACAGAGAGAAATGAAAAACCATCA 1560
 Db 1548 tcagctcgtgtaaaacagaacctctaactatacagacgaggaatgaaaaaccatcaa 1607
 OY 1561 GACCTCTCATTTGTACACCTGATGAACTATTGCTGGCTATACGGCTAATTTCAAT 1620
 Db 1608 gacctctcaatttgtagacctgtagaactattgctggtatagcgttaataat 1667
 OY 1621 GGCTGTATGAGACAGGCTATTCTTAACCTCTGACACCACTTGTAGGCAATGGCTTACGGT 1680
 Db 1668 ggcctgtatggaacagcgtatcttaacgcgtctgacacacgtgtgagcaatggccttaacggt 1727
 OY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTTCTCTGAAGAAACAATCCAGAGA 1740
 Db 1728 cgcgcgaagaagtttaaccgctctatgatacgtctgtaaggaagcaatccagagga 1787
 OY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGGAATTGCTATTAAAAATGGTCTCG 1800
 Db 1788 ttggaataatacagaagggctctacagaatgggaattcgtatttaaaatggctcgt 1847
 OY 1801 TTTCAGCTGGAACACTCTCTCTCACAAACCCCATCAACTGAAGTTCAACCTCATC 1860
 Db 1848 ttctacgttgaagctcacctcgtctcaacaacaccccccatcaactgaaagttcaagctc 1907
 OY 1861 ATCAGATAGTTCAACTCTACAGTCTAGCTCAACCATTCACAGCACAAAATATAGTACGAC 1920
 Db 1908 atcagatagttcaactctcaagttcagctcaacacacccaagaacaataatagtaagac 1967
 OY 1921 TACCAATCCATACATATATACGCAATCAAAATACAAACCCCTGATACAAATTCGAA 1980
 Db 1968 taccaatccatacaataatgcaacaatacaatacaacccctgatacaacaataatcgaa 2027
 OY 1981 TCCTCAACGACGACAAACA 1999
 Db 2028 tctcaacacgacacaaca 2046

RESULT 4
 AAH02055
 ID AAH02055 standard: DNA: 2160 BP.
 XX
 AC AAH02055:

XX 24-JUL-2001 (first entry)
 DT Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2048.
 XX
 DE
 XX Species specific: genus specific; family specific; probe: detection;
 KW identification; algal; archaeal; bacterial; fungal; parasitic;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 KW vaccine; primer: ds.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN W0200123604-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000MO-CA01150.
 XX
 PR 28-SEP-1999; 99CA-2283458.
 PR 19-MAY-2000; 2000CA-2307010.
 XX
 PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
 XX
 PI Bergeron RG, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;
 XX
 DR WPI: 2001-245006/25.
 XX
 PT Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 PT archaeal, bacterial, fungal and parasitic species in a test sample -
 XX
 PS Disclosure: Page 1463-1464; 1580pp; English.
 XX
 CC The present invention describes a method for generating a repertory of
 CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitic
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitic species, genus, family and group. A nucleic acid (I)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexa nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (I) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX
 SO Sequence 2160 BP; 697 A; 522 C; 413 G; 528 T; 0 other;

Query Match 71.9%; Score 1438; DB 22; Length 2160;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 OY 1 TAAATCTACGACATAAATCAACTATGCTGAGCTTGGGCTCTGAGCGCGCGCA 60
 Db 159 taaatctacgacataaatacaactatgctgagcttgggctctgagcgcgcgca 218

QY	61	TGCCAAGCTAAATGATATTTCCCAACAAATTTGGTTAAGGCAATCCGTTTCTATCGAAGACA	120
Db	219	tgcccaagctaaatgataatttcccacagaatttggtaaggaaatcgtttctatcgaagacca	278
QY	121	TGCGTTTGTGACACAGGGGGATTTGATACGATCCGTAATCCCGGAGGCTTTCTTGCGCAA	180
Db	279	tgcgttcttcgcacacaaggaggatgataacatccgtatcccgaggatcttcctgcgaa	338
QY	181	TCGTGCAAGCAATTTCCCTCCAGAGTGGATCAACTCTCAACCCAAAGTTGATTAAGTTGAC	240
Db	339	tcgtcaagaagaattccctcccaagtgatcagctctccactcaacagttgattaagtgtac	398
QY	241	TTACTTTTCAACTTCGCATCTCCGACCAAGCTATTTTCTGTGAAGCTCAGGAAGCTGGTT	300
Db	399	ttacttttcaacttcgactcttcgcagcacagactatcttcgttaagctcaggaagcttggtt	458
QY	301	AGCATTTCAGTTTGAACAACAAAACCAACGACAGAGAAGAAATCTTGACTCTATTAATAATA	360
Db	459	agcatttcagtttgaacaaaagaagcaaccagaagaatacttgcacctctataaataaa	518
QY	361	GGTCTACATGTCTAATGGGAATGGAATGGAATGGACAGACAGCTCAAAACTACTATGTAA	420
Db	519	ggtctacatgtctaatgggaattgggaattgggaattggacagacagctcaaaactactatgtaa	578
QY	421	AGACCTCAATATTTTAAGTTTACTTCAGTTTACCCTTGCTGGCTGGAAATGCTCTCAGGACCC	480
Db	579	agacctcaaatatTTTAAGTTTACTTCAGTTTACCCTTGCTGGCTGGAAATGCTCTCAGGACCC	638
QY	481	AAACCAATATGACCCCTATTTACATCATCAACAGACAGCCCAAGCCGCAAACTTGGCTT	540
Db	639	aaaccaaatatgacccctatttcacatccagaagcagcccaagaacgcgcaaaacttggctt	698
QY	541	ATCTGAAATGAAAAATCAAGGCTACTCTCTGCTGAACAGTATGAGAAAGACAGTCAATAC	600
Db	699	atctgaaatgaaaaatcaaggctactctctgctgaaacagtatgagaaagacagtcaaatac	758
QY	601	ACCAATTAACGATGAGACTACAAAGTCTCAATACGACAGATATTAATCCTGCTTACATGGA	660
Db	759	accaathtaactgataaggcttaacaaagctcaaatcagaatlaattcccttgcttacaatgga	818
QY	661	TAAATACCTCAAGGAAGTCAATCAATACGTTGAGAGAAACAGCTATTAACCTACTACAC	720
Db	819	taaatacctcaaggaagctcatcaatcaagttgagaagaagaaagctataacctactaac	878
QY	721	AACTGGGATGGATGTCTACACAATATTAACCAACAGACAGCTCAAAAACATCTGTGGGATAT	780
Db	879	aactgggaatgatagtctacacaatatgtagaaccaagaagctcaaaaaactcgttggatat	938
QY	781	TTACAAATACAGACGAATAGCTTGCTTCATTCACAGATGATTAATGCAAGTGCCTTCCACAT	840
Db	939	ttacaatafacagaagaaatagcttgcttaccatacgaatgaattgcaagtcgcttaccat	998
QY	841	TGTTGATGTTTCTAAGGTTAAAGTCAATTCGCCAGCTAGGAGACGCCATCACTGTAAGTAA	900
Db	999	tgttgatgtttctaaaggttaaagtcaatcttgcagctaggaagaagcctacagtaagtaa	1058
QY	901	TGTTTCCCTTGGAAATTAACCAACACAGTATGAAACAAACCGCACTGGGATCAACTATGAA	960
Db	1059	tgtttcccttcgagattaaacaagcagtagaanaacaaacgcgcacttgggagtaactatgaa	1118
QY	961	ACCGATACAGACTATGCTCTCCTCTTGAGATGAGGTGTCTACGATTCACCTGCTACTAT	1020
Db	1119	accgatcaacgaatactgctcctgccttggatagcgtgtctacgtagtcaactgcacat	1178
QY	1021	CGTTACAGATGAGACCCCTAATACTACCTCGGCAAAATCTCCTGTTTAACTCGGGATAG	1080
Db	1179	cgttcaagatgagccctataaactaaccttcggacaataatcccttgttataacttcgtagag	1238
QY	1081	GGGCTACTTTGGCAACATCACTTGGCAATACGCCCTGCAACAATGGCAAAAGTCCAGC	1140
Db	1239	gggctactttggcaacatacctccttggaataagccttcgacaaatcgcgaaagctcccaagc	1298
QY	1141	CGTGGAAACTCTAAACAGGTGCGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG	1200

Db	1299	CGTGGAAACCTCAACCAAGTCTGGACCAACCGCGCAAGACTTCCCTCAAAATGCTCGG	1358
Qy	1201	AATTCGACTCCCAACCAATATATCTACTCTCAAAATGCCATTTCAAGTAACACACGCAATTCAGA	1260
Db	1359	aatgcactaccaccaagtatctactactactaaatgccaattccaagtaacaacaaccgaaatcga	1418
Qy	1261	CAAAAAATATGAGCAAGTAGTGGAAAGATGGCTGCTCTTACGCTGCTTTGGCAATATGG	1320
Db	1419	caaaaaataatgagcaagaatgagaaagaatggctgctgcttaacgctgcttgccttgccttgc	1478
Qy	1321	TGGAACTTACTATTAACCAATGTATATTCCTATTAATCGTCTTTAGTATGGAGTGGAAAA	1380
Db	1479	tggaaacttactataaacaatgatatccataaagtcgtctttagtataaggagtgaaaa	1538
Qy	1381	AAGATTCTCTAATGTGCGGAACCTGCGCATGATGAAGAAAGCAAGACGCTTATATGATACCGA	1440
Db	1539	agaattctctaattgctcggaacctcgcccatgaaagaagaacagacgctataatgatgaccga	1598
Qy	1441	CATGATGAAGAACAGTCTTGACTTATGGAACCTGGACGAATGCGTATCTTGCTGGCTCC	1500
Db	1599	catgatagaacacagctcttgatgataagactgagcaaatgctcatcttgccttgctcc	1658
Qy	1501	TCAGCTGTGTAACAGCAACCTCTAATCTATCTACACAGAGAAATTTGAAACCCATCATGA	1560
Db	1659	tcaggcttgtaaacacaagaaaccttcaactatacagaagagaaatltgaaacacacatcaa	1718
Qy	1561	GACCTCTCAATTTGTAGACACCTGATGAACATTAATCTCTGGCTTATACGCTTAATATTCAT	1620
Db	1719	gacctctcaatttltgacgacccgtatgaaactatttgcgtgctataagcgtlaaatatcaat	1778
Qy	1621	GGCTGTATGAGCAGGCTATTTCTAACCGTCTGACACCACTTGTAGGCAATGGCCTTAGCGT	1680
Db	1779	ggctgtatgagacaggtctattctcaaccgctctgacacaccttgaaggcaatgaccttaagct	1838
Qy	1681	CGCTCCCAAGTTTACCGCTCTATGATGACCTACCTGCTCTGAAGAACCAATCCGAAAGA	1740
Db	1839	cgcctccaaaglttaacgctctatgatagacttaacctgtctgaaggaaacaaatccagagga	1898
Qy	1741	TTGGAATATATACAGAGGGGCTCTACAGAAATGGAATTCGATTTAAAAATGGTGTCTCG	1800
Db	1899	ttagaataatataccagggggctctcaagaagaatgagaaatctgatttaataaatgtygtccg	1958
Qy	1801	TTCTACGTGGGAAGCTGCTGCTCCACAAACCCCATCAACTCACTGAAAGTTCAAGCTCATC	1860
Db	1959	tcttaacgtgagctcacctgctccacaacaaccccccatcaactgaaagtccaagctatc	2018
Qy	1861	ATCAGATAGTTCAACTTACAGATCTAGACTCAACCACTCCACAGCACAATTAATAGTACAC	1920
Db	2019	atcagaatagtttaacttaactcaagctacgctacgctcaacacactccaagcaaaaataatagtagac	2078
Qy	1921	TTCCGATCCTTACATTAATATACGACAAATCAATCAATCAACCCCTGTATCAACAAAATCAGAA	1980
Db	2079	taccaatcttaacaataatagcgacaacatcaatacaaacacctgatacaacaataatcagaa	2138
Qy	1981	TCCTCAACCAAGCACAACCA 1999	
Db	2139	tcctcaacacagacaacca 2157	
RESULT	5		
AAS55688			
ID	AAS55688	standard; DNA; 2160 bp.	
AC	AAS55688;		
XX			
XX	13-FEB-2002	(first entry)	
XX			
XX	Streptococcus pneumoniae	DNA for cellular proliferation protein #259.	
KM	Antisense; ds; prokaryotic cellular proliferation gene;		
KW	antibiotic; antibacterial; drug design.		
XX			

OS Streptococcus pneumoniae.
XX
PN W0200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
DR P-PDB: AAU37829.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX
PS Claim 27; Seq ID No 9325; 511pp; English.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 2160 BP; 697 A; 522 C; 413 G; 528 T; 0 other;

Query Match 71.9%; Score 1438; DB 23; Length 2160;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTAGACAAATATAATCACTATGCTGACTGGTTCTGAGCCGCTCAA 60
DB 159 taaatctagacaataataataacatcattgctgacttgggtctgaagcgcgctcaa 218
QY 61 TGCCCAAGCTAATGATATTCCTCCACAGATTGGTTAAGCAATCGTTTCTATCGAAGCA 120
DB 219 tgcaccaagctaataatattcccaagatttggttaagcaatcgtttctatcgagaaca 278
QY 121 TCGCTTCTTGACACACAGGGGATTTGATACATCCGATCTGGGACTTTCTTGCCAA 180
DB 279 tcgcttcttgaccacacaggggattgatacatccglatcctggagcttcttgcgcaa 338
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCAACTCAACCAAGATTGATTAGTTAGAC 240
DB 339 tctgcaagaacatctccctcaagtgatgataagctctcaacaagattgattagttgac 398
QY 241 TTACTTTCACTTGCAGCTTCGACGAGACTATTTCTGTAAGGCTCAGGAAGCTTGTT 300

DB 399 ttactttcaacttcgacttcgcagcagactatttctcgttaaggtcaggaagcttgctc 458
QY 301 AGCATTCAGTTAGACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
DB 459 agcatctcagtttagaataaaagaacaaagcaagaatcttgacctataataataa 518
QY 361 GGCTACATGCTAATGGAACCTATGGAAATGACAGCAGCTCAAAATCACTATGTGTAA 420
DB 519 ggtctacatgtctaatggaactcttgaaatgacagcgcgtctaaactactctgttaa 578
QY 421 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCCTTCTGGCTGAAATGCCTCAGCACC 480
DB 579 agacctcaataatattagttacctcagttagccttgctgctgtaagtccctcagcacc 638
QY 481 AAACCAATATGACCCCTATTCACATCCAGAAAGCCCAAGACCGCGGAAACTGTGCTT 540
DB 639 aaaccaataatgacccctattcaatccagaagcagcccaagaccggaacttgctc 698
QY 541 ATCTGAATGAAAAAATCAAGGCTACATCTCTGCTGAAACAGTATGAGAAGCAGTCAATAC 600
DB 699 atctgaaatgaaaaatacaaggtctacatctctgctgaacagttatgagaagcagtaac 758
QY 601 ACCAATTAAGTGAAGCTCAAAAGCTCAAAATCAGCAAGTAAATTAACCTTGCTTACATGA 660
DB 759 accaattacagttggtctacaagctcacaatcagcaagtaattaccctgtctacatgga 818
QY 661 TAAATTAACCTCAAGGAGATCATCAAGTTGAAGAAAGAAACAGCTTAATACCTACATCAC 720
DB 819 taattaccctcaaggaagatcaatcaatcaagttgagaagaagaacaggtcatacctaccac 878
QY 721 AACTGGATGATGCTCTACACAAATGTAGACCAAGAAAGCTCAAAATCATCTGTGTGATAT 780
DB 879 aactggatgattgcttcaacaataatgtagccaagaagctcaaaaacatctgtggttat 938
QY 781 TTACAATACAGACAAATACGTTGCTTATCCAGACGATGAATTTGCAATGCTTACCAT 840
DB 939 ttaacaatacagacaagaatcagcttgcttaccagacagatgaaatcagcttctacat 998
QY 841 TGTGATGTTTCTTAACGCTTAAGCTTAAGTCCACACTGAGAGACAGCATCAAGTCAATGA 900
DB 999 tgtgattgcttcaagtgataagtaatgataatgctccagctgagagcagccatacgaagtaa 1058
QY 901 TGTTCCTTGGAATTAACCAAGCAGTAGAAGCAAAACCGGACTGGGATCAACTATGAA 960
DB 1059 tgttctcttggaattacaacaagcagtagaagaacacccgacgtgggatacactatgaa 1118
QY 961 ACCGATACAGACATATGCTCTGCTTGAGATAGGTTGTCTACGATTCACATGCTACTAT 1020
DB 1119 accgatcacagactatgctctctgcttgagatagcgtgtctacagttcaacttcgactat 1178
QY 1021 CGTTACAGATGAGCCCTATATACTACCTTGAGACAAATATCTCTTTATACTGGATAG 1080
DB 1179 cgttcaagaatgagccctataactacccctgagacaataataccctgttataacttgggtag 1238
QY 1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAATCGGAAACGTCCAGC 1140
DB 1239 gggctactttggcaacaatacacttgcaatagccctgcaacaatacggaaacgtccagc 1298
QY 1141 CGTGGAAACTTAACAAGGTGCGACTCAACCGCGCAAGACTTCTCTTAATGTGTAGG 1200
DB 1299 cgtggaactcttaacaagaatcgagatccaacgagcaagaacttctctaattgtctcgg 1358
QY 1201 AATCGACTACCAAGTATTCATCTCACTCAATTCGCAATTCCTTAAGTTAACCAACCGAATAGA 1260
DB 1359 aatcgactaccagaatattcaactctcaaatgcaattcaagtaacaacaacggaatcaga 1418
QY 1261 CAATAAATATGAGACAGTATGTAAGATGGCTGCTTACCTGCTTGAATAATGG 1320
DB 1419 caataaataatgagacagatgtaagatggctgcttaccctgcttgaataatgg 1478
QY 1321 TGAAGTTACTATTAACCAATGATATATCAATAAAGTGTCTTTAGTATGGAGATGAAA 1380

Db	1479	tggaacttactataaaccaatgylatalccataaagtcgylctttagtgbtggagtgtaaa	1538
Qy	1381	AGAGCTCTTAATGTGCGAACTCGTGCCTTGAAGAAGACAGCCCTATATGATGACCGA	1440
Db	1539	agagttctctaattgctggaactcgtgcataagaaagacagcgtatattgatgacgga	1598
Qy	1441	CATGATGAAAAACAGTCTTTGACTTATATGGAACTGGACAGCAAAATGCCCTATCTTGGCTCC	1500
Db	1599	catgatbaaaacagcttctgtagtattggaactggagcaaatgacctactgtcttgcgtcc	1658
Qy	1501	TCAGGCTGCTAAAACAGAGCACTCTTAACATATACAGACGAGAAATTTGAAACCAATCAT	1560
Db	1659	tcagagctcggtaaaacaggaaaccttaactatacaagcagaggaatttgaaacacacaa	1718
Qy	1561	GACCTCTCAATTTGTATGACACCTGATGAACTATTTGGCTGTATACGCTTAATATTCAT	1620
Db	1719	gaccttcaactttgtagcacttgatgaactattgtgtctatacgcglaaatatcaat	1778
Qy	1621	GGCTGTATGACAGGCTATTTCTTACCGCTGTGACACACCTTTGAGCAATGGCTTACGCT	1680
Db	1779	ggctgtatgacagggctattcttaaccgcttgacacactttaggcaatgyccttaagct	1838
Qy	1681	CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCGTGTGAAGGACCAATCCAGAA	1740
Db	1839	cgcgcgcaaaagtttaaccgctctatgtagaccttaacctgtctgaaggaaacatccaagga	1898
Qy	1741	TTGGAATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTATTTAAAAATGGTCTCG	1800
Db	1899	ttggaatatataccagggggctcttaacgaaatggagaattcgatttaaaatgtygtccg	1958
Qy	1801	TTTACGTGGAACCTCACCTGCTCTCCACACACCCCCATCACTGAATGTAAGCTATC	1860
Db	1959	tttaacgtggagctcactcgtctcaacaacaccccatcaactgaagttaacggtc	2018
Qy	1861	ATCGATATGTTCAACTTCACAGCTGTAGTGCCTAACCAATCCAAATAATATAGTACAC	1920
Db	2019	atcgaataagttcaactctcaagctcgtactcaaccactccaagcacaataatagtagac	2078
Qy	1921	TACCAATCTCAAAATATATACGCAACAATCAATAATCAACCCCTGATCAACAAATCAGA	1980
Db	2079	taccaatctcaataataatagcgaacaataccaataacccctgtatcaacaataatcgaa	2138
Qy	1981	TCCTCAACGACGACACCA 1999	
Db	2139	tcctcaaccagcacaacca 2157	
RESULT 6			
AAH02067	ID	AAH02067 standard; DNA; 1249 BP.	
XX	XX	AAH02067;	
XX	XX	24-JUL-2001 (first entry)	
DE	DE	Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2060.	
XX	XX		
OS	OS	Streptococcus pneumoniae.	
XX	XX		
PN	PN	MO200123604-A2.	
XX	XX		
PD	PD	05-APR-2001.	
XX	XX		
PF	PF	28-SEP-2000; 2000MO-CA01150.	
XX	XX		
PR	PR	28-SEP-1999; 99CA-2283458.	
<p>Species specific; genus specific; family specific; probe; detection; Identification; algal; archaeal; bacterial; fungal; parasitical; microorganism; diagnosis; translation elongation factor Tu; toxin; translational elongation factor G; RecA recombinase; resistance; catalytic subunit of proton-translocating ATPase; antimicrobial; vaccine; primer; ds.</p>			

PR	19-MAY-2000; 2000CA-2307010.
XX	(INFE-) INFECTION DIAGNOSTIC (IDI) INC.
PA	Bergeron MG, Boleslnot M, Huletsky A, Menard C, Ouellette M;
XX	Picard FJ, Roy PH;
PI	WPI: 2001-245006/25.
XX	
DR	Nucleic acid sequences are used to generate universal probes and
XX	primers which can be used to identify and detect the presence of algal,
PT	archaeal, bacterial, fungal and parasitital species in a test sample -
XX	
PS	Claim 27; Page 1473; 1580pp; English.
XX	
CC	The present invention describes a method for generating a repertory of
CC	nucleic acids of tuf, fns, atp and/or recA genes from which probes
CC	and/or primers are derived. The method comprises amplifying the nucleic
CC	acids of determined algal, archaeal, bacterial, fungal and parasitcal
CC	species with a combination of defined primer pairs. The method can be
CC	used for producing probes and/or primers for detecting one or more
CC	related microorganisms e.g. algae, archaea, bacteria, fungi and
CC	parasites, for universal detection and for specific and ubiquitous
CC	detection and identification of an algal, archaeal, bacterial, fungal
CC	and parasitcal species, genus, family and group. A nucleic acid (1)
CC	obtained using the method of the invention can be used for the universal
CC	detection of any bacterium, fungus or parasite in a sample and for the
CC	detection of at least one antimicrobial agent resistance gene or at
CC	least one toxin gene. hexa nucleic acids are used for the specific and
CC	ubiquitous detection and for identification of Streptococcus pneumoniae.
CC	(1) can be used to design a therapeutic agent which is effective against
CC	microorganisms. Microbial species or genus or family or phylum or group
CC	which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC	Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC	Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC	Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC	provides faster results than substrate specificity tests as results can
CC	be determined in an hour and improved accuracy is also achieved.
CC	AHH00010 to AHH002304 represent nucleotide sequences and primers/probes
CC	which are given in the exemplification of the present invention.
XX	
XX	
SQ	Sequence 1249 BP; 402 A; 302 C; 246 G; 299 T; 0 other:
	Query Match 59.9%; Score 1198; DB 22; Length 1249;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 1248; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	720 CAACGTGGATGATGTCTACACAAATGTAGACAGAAGCCTCAAAAACATCTGTGGGATA 779
Dd	1 caactcggatgatgtctaccacaattgtagaccaagaagctcaaaaacctctgysgata 60
OY	780 TTTCACAATACAGACGAATACGTTGCCATTCCAGACGATGATTCGAAGTGCCTTCTACCA 839
Dd	61 tttaacaatacagaagaatcgltgcctaccagaaatgaattgcgaagtcgcttccacca 120
OY	840 TTGTTGATGTTTCTAACGGTAAAGCATTTGGCCACCAGTGGACACAGCCACTCGTAGTAAGTA 899
Dd	121 ttgttgatgattttctaaggtaaagtcattgccagcttagagcacagccaatcaglaagta 180
OY	900 ATGTTTCCCTTGGAATTAACCAAGCATGTAAGAAACAACCGCACCTGGGANTCAACTATGA 959
Dd	181 atgttcccttggaaattaaccagaagcatagaanaaacacgcgactggggataactaataya 240
Dd	241 aaccgatcacagaactatgctcctgccttgagtaacgltgcttaacatcaactgctacta 300
OY	1020 TCGTTCACAGATGAGCCCTATACCTACCCCTGGGCAAAATACCTCTTTATATCGGGATA 1079
Dd	301 tcgttcacagatgagccctataactaccctgggacaataactcctglttataactaygata 360
OY	1080 GGGGGTACTTTGGCAACAATCACCATTGCAATAGCCCTCAACAAATCGGAAACGTCCAG 1139

Db 361 | 999ggaactcttggcaacatcaccttgcatacagccctgcacacatcgaaacgtccag 420
Oy 1140 | CCGTGAAGAACTCTAAACAAAGGTGGAGCTCAACCGCGCAAGACTTTCCTAAATGCTAG 1199
Db 421 | cgtggaactctaaacaaggtcgcgaactcaacgcgcgaagacttccctaaatggtctag 480
Oy 1200 | GAATGACATACCCAGATGATCTACTACTCAATGCAATTCCTCAAGTAAACACAGCAATCAG 1259
Db 481 | gaatgcatacccaagatattcactactcaaatgcattcaagtaacacacgaatcag 540
Oy 1260 | ACAAATAATATGAGCAAGTAGGAAAGATGGCTGCTGCTTACGCTTGCTTGCAATG 1319
Db 541 | acaaaaaaatgagcaagtagyaaaagatggctgcttcaacgcttgccttgcgaatg 600
Oy 1320 | GTGGAAGTCTACTTAACCAATGTATATCCATTAAGTCTGCTTACGATGAGGATGA 1379
Db 601 | gttggaacttactataaacaatgataatcaataaagtcgtctttagtgatggagtgaaa 660
Oy 1380 | AAGAGTTCCTAAATGTCGGAAGTCTGTCATGAGGAAGAGAGAGCTATATGATGACCG 1439
Db 661 | aagagttctctaattgctgcgaactcgtgcataaggaagaaagacagctatagatgacgg 720
Oy 1440 | AATATGATGAAGAACTGCTGACTTATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1499
Db 721 | aatgtgtaaaacagctctgacttgaactgagacgaatgtcctatcttgcctgctcc 780
Oy 1500 | CTCGAGCTGTTAAACAGAGAACCTTAACTATACAGACGAGAAATGTAACACATCA 1559
Db 781 | ctcagagctgtgtaaaacaggaactcctactatacagaggaagaaatgtgaaaccacatca 840
Oy 1560 | AGACCTCTCAATTTGAGACACCTGATGAACTATTGCTGCTATACCGCTAATATTCA 1619
Db 841 | agacctctcaatttgaagacccgtgaaactatttgcgtctatacgcgttaaatattca 900
Oy 1620 | TGGCTTATGAGACAGGCTATTCTTAACGCTGTGACACACTGTTAGCAATGGCTTACGG 1679
Db 901 | tggcttataaggaagacttacttaacgcgtgcgacacacttgaagcaaatggccttaagg 960
Oy 1680 | TCGCTGCAAAAGTTTACCGCTCTATGATGACCTACTGCTGTAAGAGAGCAATCCAGAA 1739
Db 961 | tcgctgcacaaagttaacgctctatgatgacactcgtctggaaggaatccagaag 1020
Oy 1740 | ATTGGAATATACAGAGGGGCTCTACAGAAATGGAATTCGATTATTAATATGCTGCTC 1799
Db 1021 | attggaatattaccagaaggggtcttacagaatlgagaatctgatttaaaatgggtgctc 1080
Oy 1800 | GTTCTACGAGAACCTACCGTGTCCACAAACCCCATCAATGGAAGTTCAAGGTCAT 1859
Db 1081 | gtctcgtggaagctcaactgctccacaacaccccaacacacacacacacacacacacacac 1140
Oy 1860 | CATCAGATAGTTCAACTTACAGTCTAGCTCAACACTCCAAACACACAAATATATGATGCA 1919
Db 1141 | catcagatagttcaacttaccagcttagctcaacacacacacacacacacacacacacacac 1200
Oy 1920 | CTACCAATCTTAACATATATAGCAACATCAATCAACCCCTGATCA 1968
Db 1201 | ctaccaatcttaacaataataacgcaacaatacaatacaacccctgta 1249

RESULT 7
AAZ35952 standard: DNA; 1260 BP.
XX AAZ35952;
XX
XX 07-FEB-2000 (first entry)
XX
DE Streptococcus pneumoniae pbp1A TER isolate n) nucleotide sequence.
XX
XX Streptococcus pneumoniae; penicillin binding protein: pbp2B; pbp1A;
XX transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
KM detection; identification; pneumococcal meningitis; ss.

XX OS Streptococcus pneumoniae.
XX PN ZA9807024-A.
XX XX
PD 28-APR-1999.
XX
PF 05-AUG-1998; 982A-0007024.
XX
PR 01-AUG-1997; 972A-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UYWI-) UNIV WITWATERSRAND.
PA (MED-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
PI WPI; 1999-601770/51.
DR
XX
XX Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis
XX
PS Claim 11; Fig 4; 63pp: English.
XX
CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1A gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics. Preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1A transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX
SQ Sequence 1260 BP; 407 A; 292 C; 259 G; 302 T; 0 other;

Query Match 50.7%; Score 1014; DB 20; Length 1260;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1214; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 617 CTACCAAGTCTCAATCAGCAAGTATTTACCTGCTTACATGATGATTAATCTCAAGAA 676
Db 43 ctacaaagttcaaatcagcaagtaattacccgtcttacaatgataattaccctcaaggaa 102
Oy 677 GTCATCAATCAAGTGTGAAGAAGAACAGGCTATACCTCACTCAACTGGATGATGTC 736
Db 103 gtcatcaatcaagttgaaagaagaacaggtctataactctcaacactggtgagtgtc 162
Oy 737 TACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATATTTACATACAGACGA 796
Db 163 tacacaaatgtagacccaagaagctcaaaaacatctgtggatatttacaatacagcaga 222
Oy 797 TACGTGCTATTCACAGATGATGATGCAAGTGTGCTTACCAATGTTGATGTTTAAAC 856
Db 223 taagttgctctatccacagatgatatgcaagtgctctaccatctgtgatgttttcaac 282
Oy 857 GGTAAAGTCAATGCCCAGGTAGAGACCGCATCAGTCAAGTAATGTTTCTCGGAATT 916
Db 283 ggtaaagtcatgcccagctagagacagccatcaagtaaatgttccctcggaaatt 342
Oy 917 AACCAAGCAGTAGAACAACCGCACTGGGATCAACTATGAAACCGATCACAGACTAT 976
Db 343 aaccaagcagtagaacaacacgcgactggtgatacactatgaaacacgatacaagctat 402
Oy 977 GCTCTGCTTGAGAGTACGGTGTACAGATTCAACTGCTACTACTGTTACAGATGAGCCC 1036


```

Db 403 gctccgcttagtagagagtggtctacgaltcaactgctacatcgltcacgagtagccc 462
Oy 1037 TATTAATACCCGGGAGCAATACCTGCTTTAATCGGATAGGGGCTACTTTGGCAAC 1096
Db 463 tataatacccccggagcaaatactccgtcttaacttggaataagggtcactttggcaac 522
Oy 1097 ATACACCTTGCAATAGACCCCTGCAACATCGCAAAAGTCCAGCCGTGGAACCTGTAAC 1156
Db 523 atcaccttgcaatacgccttcgcaacaatcgcaaacgtcccgctggaaactctaac 582
Oy 1157 AAGGTGGAGCTAACCCGGCCAGACTTTCTTAATGCTTAGGAATGACTACCCAACT 1216
Db 583 aagtcgagctaacacgcgcgaagcttccctaagtgtctaggaatcgactaccaag 642
Oy 1217 ATTCACTACTCAAAATGCCCTTTCAGTATACACACCGAATCAGACAAAATATGAGACA 1276
Db 643 attcaactactaaatgaccttcaagtaacacaccgaatcgacaaaaatattgagca 702
Oy 1277 ACTAGTGAAGATGGCTGCTGCTTACGCTGCTTGGCAATGGTGAAGTACTATATA 1336
Db 703 agtagtgaagaagatggtcgtctgctacgtcgtccttgcaaatggtggaactactataa 762
Oy 1337 CCAATGTATATCCATAAAGTGTCTTATGATGGAGTGAAGAAAGTCTCTAATGTC 1396
Db 763 ccaatgtatatccataaagtcgtcttagtgatggagtgagaaagagttctctaatgctc 822
Oy 1397 GGAACCTCGCCGATAGAGAAAGACAGACGCTTATGATATACCGACATGATGAAAACAGTC 1456
Db 823 ggaacctcgccatgaagaaagacagcctatatgatagacccgacatgtagtaaacagtc 882
Oy 1457 TTGACTTATGAACTGAGAAATGCTATCTTGGCTTGCTGCTCCCTCAGGCTGTAAACA 1516
Db 883 ttgagttatggaactcggaacgaatgctctatctgcttgctccctcagctgtgtaaaaa 942
Oy 1517 GGAACCTCTAATATATACAGACAGAAATGTAACACACATCAAGCTCTCAATTGTA 1576
Db 943 ggaacctctaactatacagacgaggaatgtaaaacacatcaagaacctctcaattgt 1002
Oy 1577 GCACCTGATGAACTATTTGCTGCTATACGCGTAATATTCATATGCTATGACGAGC 1636
Db 1003 gcaactgaagactattgctggtcctatacgcgttaaatatcaatgctgtaaggcaagc 1062
Oy 1637 TATTTAAACGCTGTGACACCACTTGTAGCAATGGCTTACGCTGCTGCCAAAGTTTAC 1696
Db 1063 tatttaacccgtctgacacccacttgtagcaatggtccttaagctggtccaaagttaac 1122
Oy 1697 CGCTCTATGATACCTACTGCTGTGAAGAGCAATCCAGAAGATTGGATTTACCAAG 1756
Db 1123 cgcctctatgatacctactgtctgaaggaacatccagagagattggaatataccagag 1182
Oy 1757 GGGCTCTACAGAAATGAGAAATTCGATTAAATGGTGGTCTTACGCTGAGAACTCA 1816
Db 1183 gggcctctacagaatggaatctgcttaataaatggtccgttcttaagctggaactca 1242
Oy 1817 CCTGCTCCACAACAACCC 1834
Db 1243 cctgctccaacaacacc 1260

```

RESULT 8
 AA235953
 ID AA235953 standard; DNA; 1260 BP.
 AC AA235953;
 DI 07-FEB-2000 (first entry)
 DE Streptococcus pneumoniae pbp1A T8R isolate o) nucleotide sequence.
 XX Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;
 KM transpeptidase encoding region; T8R; antibiotic resistance; diagnosis;
 KM detection; Identification; pneumococcal meningitis; ss.

```

XX 05 Streptococcus pneumoniae.
XX 06
XX 07
XX 08
XX 09
XX 10
XX 11
XX 12
XX 13
XX 14
XX 15
XX 16
XX 17
XX 18
XX 19
XX 20
XX 21
XX 22
XX 23
XX 24
XX 25
XX 26
XX 27
XX 28
XX 29
XX 30
XX 31
XX 32
XX 33
XX 34
XX 35
XX 36
XX 37
XX 38
XX 39
XX 40
XX 41
XX 42
XX 43
XX 44
XX 45
XX 46
XX 47
XX 48
XX 49
XX 50
XX 51
XX 52
XX 53
XX 54
XX 55
XX 56
XX 57
XX 58
XX 59
XX 60
XX 61
XX 62
XX 63
XX 64
XX 65
XX 66
XX 67
XX 68
XX 69
XX 70
XX 71
XX 72
XX 73
XX 74
XX 75
XX 76
XX 77
XX 78
XX 79
XX 80
XX 81
XX 82
XX 83
XX 84
XX 85
XX 86
XX 87
XX 88
XX 89
XX 90
XX 91
XX 92
XX 93
XX 94
XX 95
XX 96
XX 97
XX 98
XX 99
XX 100
XX 101
XX 102
XX 103
XX 104
XX 105
XX 106
XX 107
XX 108
XX 109
XX 110
XX 111
XX 112
XX 113
XX 114
XX 115
XX 116
XX 117
XX 118
XX 119
XX 120
XX 121
XX 122
XX 123
XX 124
XX 125
XX 126
XX 127
XX 128
XX 129
XX 130
XX 131
XX 132
XX 133
XX 134
XX 135
XX 136
XX 137
XX 138
XX 139
XX 140
XX 141
XX 142
XX 143
XX 144
XX 145
XX 146
XX 147
XX 148
XX 149
XX 150
XX 151
XX 152
XX 153
XX 154
XX 155
XX 156
XX 157
XX 158
XX 159
XX 160
XX 161
XX 162
XX 163
XX 164
XX 165
XX 166
XX 167
XX 168
XX 169
XX 170
XX 171
XX 172
XX 173
XX 174
XX 175
XX 176
XX 177
XX 178
XX 179
XX 180
XX 181
XX 182
XX 183
XX 184
XX 185
XX 186
XX 187
XX 188
XX 189
XX 190
XX 191
XX 192
XX 193
XX 194
XX 195
XX 196
XX 197
XX 198
XX 199
XX 200
XX 201
XX 202
XX 203
XX 204
XX 205
XX 206
XX 207
XX 208
XX 209
XX 210
XX 211
XX 212
XX 213
XX 214
XX 215
XX 216
XX 217
XX 218
XX 219
XX 220
XX 221
XX 222
XX 223
XX 224
XX 225
XX 226
XX 227
XX 228
XX 229
XX 230
XX 231
XX 232
XX 233
XX 234
XX 235
XX 236
XX 237
XX 238
XX 239
XX 240
XX 241
XX 242
XX 243
XX 244
XX 245
XX 246
XX 247
XX 248
XX 249
XX 250
XX 251
XX 252
XX 253
XX 254
XX 255
XX 256
XX 257
XX 258
XX 259
XX 260
XX 261
XX 262
XX 263
XX 264
XX 265
XX 266
XX 267
XX 268
XX 269
XX 270
XX 271
XX 272
XX 273
XX 274
XX 275
XX 276
XX 277
XX 278
XX 279
XX 280
XX 281
XX 282
XX 283
XX 284
XX 285
XX 286
XX 287
XX 288
XX 289
XX 290
XX 291
XX 292
XX 293
XX 294
XX 295
XX 296
XX 297
XX 298
XX 299
XX 300
XX 301
XX 302
XX 303
XX 304
XX 305
XX 306
XX 307
XX 308
XX 309
XX 310
XX 311
XX 312
XX 313
XX 314
XX 315
XX 316
XX 317
XX 318
XX 319
XX 320
XX 321
XX 322
XX 323
XX 324
XX 325
XX 326
XX 327
XX 328
XX 329
XX 330
XX 331
XX 332
XX 333
XX 334
XX 335
XX 336
XX 337
XX 338
XX 339
XX 340
XX 341
XX 342
XX 343
XX 344
XX 345
XX 346
XX 347
XX 348
XX 349
XX 350
XX 351
XX 352
XX 353
XX 354
XX 355
XX 356
XX 357
XX 358
XX 359
XX 360
XX 361
XX 362
XX 363
XX 364
XX 365
XX 366
XX 367
XX 368
XX 369
XX 370
XX 371
XX 372
XX 373
XX 374
XX 375
XX 376
XX 377
XX 378
XX 379
XX 380
XX 381
XX 382
XX 383
XX 384
XX 385
XX 386
XX 387
XX 388
XX 389
XX 390
XX 391
XX 392
XX 393
XX 394
XX 395
XX 396
XX 397
XX 398
XX 399
XX 400
XX 401
XX 402
XX 403
XX 404
XX 405
XX 406
XX 407
XX 408
XX 409
XX 410
XX 411
XX 412
XX 413
XX 414
XX 415
XX 416
XX 417
XX 418
XX 419
XX 420
XX 421
XX 422
XX 423
XX 424
XX 425
XX 426
XX 427
XX 428
XX 429
XX 430
XX 431
XX 432
XX 433
XX 434
XX 435
XX 436
XX 437
XX 438
XX 439
XX 440
XX 441
XX 442
XX 443
XX 444
XX 445
XX 446
XX 447
XX 448
XX 449
XX 450
XX 451
XX 452
XX 453
XX 454
XX 455
XX 456
XX 457
XX 458
XX 459
XX 460
XX 461
XX 462
XX 463
XX 464
XX 465
XX 466
XX 467
XX 468
XX 469
XX 470
XX 471
XX 472
XX 473
XX 474
XX 475
XX 476
XX 477
XX 478
XX 479
XX 480
XX 481
XX 482
XX 483
XX 484
XX 485
XX 486
XX 487
XX 488
XX 489
XX 490
XX 491
XX 492
XX 493
XX 494
XX 495
XX 496
XX 497
XX 498
XX 499
XX 500
XX 501
XX 502
XX 503
XX 504
XX 505
XX 506
XX 507
XX 508
XX 509
XX 510
XX 511
XX 512
XX 513
XX 514
XX 515
XX 516
XX 517
XX 518
XX 519
XX 520
XX 521
XX 522
XX 523
XX 524
XX 525
XX 526
XX 527
XX 528
XX 529
XX 530
XX 531
XX 532
XX 533
XX 534
XX 535
XX 536
XX 537
XX 538
XX 539
XX 540
XX 541
XX 542
XX 543
XX 544
XX 545
XX 546
XX 547
XX 548
XX 549
XX 550
XX 551
XX 552
XX 553
XX 554
XX 555
XX 556
XX 557
XX 558
XX 559
XX 560
XX 561
XX 562
XX 563
XX 564
XX 565
XX 566
XX 567
XX 568
XX 569
XX 570
XX 571
XX 572
XX 573
XX 574
XX 575
XX 576
XX 577
XX 578
XX 579
XX 580
XX 581
XX 582
XX 583
XX 584
XX 585
XX 586
XX 587
XX 588
XX 589
XX 590
XX 591
XX 592
XX 593
XX 594
XX 595
XX 596
XX 597
XX 598
XX 599
XX 600
XX 601
XX 602
XX 603
XX 604
XX 605
XX 606
XX 607
XX 608
XX 609
XX 610
XX 611
XX 612
XX 613
XX 614
XX 615
XX 616
XX 617
XX 618
XX 619
XX 620
XX 621
XX 622
XX 623
XX 624
XX 625
XX 626
XX 627
XX 628
XX 629
XX 630
XX 631
XX 632
XX 633
XX 634
XX 635
XX 636
XX 637
XX 638
XX 639
XX 640
XX 641
XX 642
XX 643
XX 644
XX 645
XX 646
XX 647
XX 648
XX 649
XX 650
XX 651
XX 652
XX 653
XX 654
XX 655
XX 656
XX 657
XX 658
XX 659
XX 660
XX 661
XX 662
XX 663
XX 664
XX 665
XX 666
XX 667
XX 668
XX 669
XX 670
XX 671
XX 672
XX 673
XX 674
XX 675
XX 676
XX 677
XX 678
XX 679
XX 680
XX 681
XX 682
XX 683
XX 684
XX 685
XX 686
XX 687
XX 688
XX 689
XX 690
XX 691
XX 692
XX 693
XX 694
XX 695
XX 696
XX 697
XX 698
XX 699
XX 700
XX 701
XX 702
XX 703
XX 704
XX 705
XX 706
XX 707
XX 708
XX 709
XX 710
XX 711
XX 712
XX 713
XX 714
XX 715
XX 716
XX 717
XX 718
XX 719
XX 720
XX 721
XX 722
XX 723
XX 724
XX 725
XX 726
XX 727
XX 728
XX 729
XX 730
XX 731
XX 732
XX 733
XX 734
XX 735
XX 736
XX 737
XX 738
XX 739
XX 740
XX 741
XX 742
XX 743
XX 744
XX 745
XX 746
XX 747
XX 748
XX 749
XX 750
XX 751
XX 752
XX 753
XX 754
XX 755
XX 756
XX 757
XX 758
XX 759
XX 760
XX 761
XX 762
XX 763
XX 764
XX 765
XX 766
XX 767
XX 768
XX 769
XX 770
XX 771
XX 772
XX 773
XX 774
XX 775
XX 776
XX 777
XX 778
XX 779
XX 780
XX 781
XX 782
XX 783
XX 784
XX 785
XX 786
XX 787
XX 788
XX 789
XX 790
XX 791
XX 792
XX 793
XX 794
XX 795
XX 796
XX 797
XX 798
XX 799
XX 800
XX 801
XX 802
XX 803
XX 804
XX 805
XX 806
XX 807
XX 808
XX 809
XX 810
XX 811
XX 812
XX 813
XX 814
XX 815
XX 816
XX 817
XX 818
XX 819
XX 820
XX 821
XX 822
XX 823
XX 824
XX 825
XX 826
XX 827
XX 828
XX 829
XX 830
XX 831
XX 832
XX 833
XX 834
XX 835
XX 836
XX 837
XX 838
XX 839
XX 840
XX 841
XX 842
XX 843
XX 844
XX 845
XX 846
XX 847
XX 848
XX 849
XX 850
XX 851
XX 852
XX 853
XX 854
XX 855
XX 856
XX 857
XX 858
XX 859
XX 860
XX 861
XX 862
XX 863
XX 864
XX 865
XX 866
XX 867
XX 868
XX 869
XX 870
XX 871
XX 872
XX 873
XX 874
XX 875
XX 876
XX 877
XX 878
XX 879
XX 880
XX 881
XX 882
XX 883
XX 884
XX 885
XX 886
XX 887
XX 888
XX 889
XX 890
XX 891
XX 892
XX 893
XX 894
XX 895
XX 896
XX 897
XX 898
XX 899
XX 900
XX 901
XX 902
XX 903
XX 904
XX 905
XX 906
XX 907
XX 908
XX 909
XX 910
XX 911
XX 912
XX 913
XX 914
XX 915
XX 916
XX 917
XX 918
XX 919
XX 920
XX 921
XX 922
XX 923
XX 924
XX 925
XX 926
XX 927
XX 928
XX 929
XX 930
XX 931
XX 932
XX 933
XX 934
XX 935
XX 936
XX 937
XX 938
XX 939
XX 940
XX 941
XX 942
XX 943
XX 944
XX 945
XX 946
XX 947
XX 948
XX 949
XX 950
XX 951
XX 952
XX 953
XX 954
XX 955
XX 956
XX 957
XX 958
XX 959
XX 960
XX 961
XX 962
XX 963
XX 964
XX 965
XX 966
XX 967
XX 968
XX 969
XX 970
XX 971
XX 972
XX 973
XX 974
XX 975
XX 976
XX 977
XX 978
XX 979
XX 980
XX 981
XX 982
XX 983
XX 984
XX 985
XX 986
XX 987
XX 988
XX 989
XX 990
XX 991
XX 992
XX 993
XX 994
XX 995
XX 996
XX 997
XX 998
XX 999
XX 1000
XX 1001
XX 1002
XX 1003
XX 1004
XX 1005
XX 1006
XX 1007
XX 1008
XX 1009
XX 1010
XX 1011
XX 1012
XX 1013
XX 1014
XX 1015
XX 1016
XX 1017
XX 1018
XX 1019
XX 1020
XX 1021
XX 1022
XX 1023
XX 1024
XX 1025
XX 1026
XX 1027
XX 1028
XX 1029
XX 1030
XX 1031
XX 1032
XX 1033
XX 1034
XX 1035
XX 1036
XX 1037
XX 1038
XX 1039
XX 1040
XX 1041
XX 1042
XX 1043
XX 1044
XX 1045
XX 1046
XX 1047
XX 1048
XX 1049
XX 1050
XX 1051
XX 1052
XX 1053
XX 1054
XX 1055
XX 1056
XX 1057
XX 1058
XX 1059
XX 1060
XX 1061
XX 1062
XX 1063
XX 1064
XX 1065
XX 1066
XX 1067
XX 1068
XX 1069
XX 1070
XX 1071
XX 1072
XX 1073
XX 1074
XX 1075
XX 1076
XX 1077
XX 1078
XX 1079
XX 1080
XX 1081
XX 1082
XX 1083
XX 1084
XX 1085
XX 1086
XX 1087
XX 1088
XX 1089
XX 1090
XX 1091
XX 1092
XX 1093
XX 1094
XX 1095
XX 1096
XX 1097
XX 1098
XX 1099
XX 1100
XX 1101
XX 1102
XX 1103
XX 1104
XX 1105
XX 1106
XX 1107
XX 1108
XX 1109
XX 1110
XX 1111
XX 1112
XX 1113
XX 1114
XX 1115
XX 1116
XX 1117
XX 1118
XX 1119
XX 1120
XX 1121
XX 1122
XX 1123
XX 1124
XX 1125
XX 1126
XX 1127
XX 1128
XX 1129
XX 1130
XX 1131
XX 1132
XX 1133
XX 1134
XX 1135
XX 1136
XX 1137
XX 1138
XX 1139
XX 1140
XX 1141
XX 1142
XX 1143
XX 1144
XX 1145
XX 1146
XX 1147
XX 1148
XX 1149
XX 1150
XX 1151
XX 1152
XX 1153
XX 1154
XX 1155
XX 1156
XX 1157
XX 1158
XX 1159
XX 1160
XX 1161
XX 1162
XX 1163
XX 1164
XX 1165
XX 1166
XX 1167
XX 1168
XX 1169
XX 1170
XX 1171
XX 1172
XX 1173
XX 1174
XX 1175
XX 1176
XX 1177
XX 1178
XX 1179
XX 1180
XX 1181
XX 1182
XX 1183
XX 1184
XX 1185
XX 1186
XX 1187
XX 1188
XX 1189
XX 1190
XX 1191
XX 1192
XX 1193
XX 1194
XX 1195
XX 1196
XX 1197
XX 1198
XX 1199
XX 1200
XX 1201
XX 1202
XX 1203
XX 1204
XX 1205
XX 1206
XX 1207
XX 1208
XX 1209
XX 1210
XX 1211
XX 1212
XX 1213
XX 1214
XX 1215
XX 1216
XX 1217
XX 1218
XX 1219
XX 1220
XX 1221
XX 1222
XX 1223
XX 1224
XX 1225
XX 1226
XX 1227
XX 1228
XX 1229
XX 1230
XX 1231
XX 1232
XX 1233
XX 1234
XX 1235
XX 1236
XX 1237
XX 1238
XX 1239
XX 1240
XX 1241
XX 1242
XX 1243
XX 1244
XX 1245
XX 1246
XX 1247
XX 1248
XX 1249
XX 1250
XX 1251
XX 1252
XX 1253
XX 1254
XX 1255
XX 1256
XX 1257
XX 1258
XX 1259
XX 1260
XX 1261
XX 1262
XX 1263
XX 1264
XX 1265
XX 1266
XX 1267
XX 1268
XX 1269
XX 1270
XX 1271
XX 1272
XX 1273
XX 1274
XX 1275
XX 1276
XX 1277
XX 1278
XX 1279
XX 1280
XX 1281
XX 1282
XX 1283
XX 1284
XX 1285
XX 1286
XX 1287
XX 1288
XX 1289
XX 1290
XX 1291
XX 1292
XX 1293
XX 1294
XX 1295
XX 1296
XX 1297
XX 1298
XX 1299
XX 1300
XX 1301
XX 1302
XX 1303
XX 1304
XX 1305
XX 1306
XX 1307
XX 1308
XX 1309
XX 1310
XX 1311
XX 1312
XX 1313
XX 1314
XX 1315
XX 1316
XX 1317
XX 1318
XX 1319
XX 1320
XX 1321
XX 1322
XX 1323
XX 1324
XX 1325
XX 1326
XX 1327
XX 1328
XX 1329
XX 1330
XX 1331
XX 1332
XX 1333
XX 1334
XX 1335
XX 1336
XX 1337
XX 1338
XX 1339
XX 1340
XX 1341
XX 1342
XX 1343
XX 1344
XX 1345
XX 1346
XX 1347
XX 1348
XX 1349
XX 1350
XX 1351
XX 1352
XX 1353
XX 1354
XX 1355
XX 1356
XX 1357
XX 1358
XX 1359
XX 1360
XX 1361
XX 1362
XX 1363
XX 1364
XX 1365
XX 1366
XX 1367
XX 1368
XX 1369
XX 1370
XX 1371
XX 1372
XX 1373
XX 1374
XX 1375
XX 1376
XX 1377
XX 1378
XX 1379
XX 1380
XX 1381
XX 1382
XX 1383
XX 1384
XX 1385
XX 1386
XX 1387
XX 1388
XX 1389
XX 1390
XX 1391
XX 1392
XX 1393
XX 1394
XX 1395
XX 1396
XX 1397
XX 1398
XX 1399
XX 1400
XX 1401
XX 1402
XX 1403
XX 1404
XX 1405
XX 1406
XX 1407
XX 1408
XX 1409
XX 1410
XX 1411
XX 1412
XX 1413
XX 1414
XX 1415
XX 1416
XX 1417
XX 1418
XX 1419
XX 1420
XX 1421
XX 1422
XX 1423
XX 1424
XX 1425
XX 1426
XX 1427
XX 1428
XX 1429
XX 1430
XX 1431
XX 1432
XX 1433
XX 1434
XX 1435
XX 1436
XX 1437
XX 1438
XX 1439
XX 1440
XX 1441
XX 1442
XX 1443
XX 1444
XX 1445
XX 1446
XX 1447
XX 1448
XX 1449
XX 1450
XX 1451
XX 1452
XX 1453
XX 1454
XX 1455
XX 1456
XX 1457
XX 1458
XX 1459
XX 1460
XX 1461
XX 1462
XX 1463
XX 1464
XX 1465
XX 1466
XX 1467
XX 1468
XX 1469
XX 1470
XX 1471
XX 1472
XX 1473
XX 1474
XX 1475
XX 1476
XX 1477
XX 1478
XX 1479
XX 1480
XX 1481
XX 1482
XX 1483
XX 1484
XX 1485
XX 1486
XX 1487
XX 1488
XX 1489
XX 1490
XX 1491
XX 1492
XX 1493
XX 1494
XX 1495
XX 1496
XX 1497
XX 1498
XX 1499
XX 1500
XX 1501
XX 1502
XX 1503
XX 1504
XX 1505
XX 1506
XX 1507
XX 1508
XX 1509
XX 1510
XX 1511
XX 1512
XX 1513
XX 1514
XX 1515
XX 1516
XX 1517
XX 1518
XX 1519
XX 1520
XX 1521
XX 1522
XX 1523
XX 1524
XX 1525
XX 1526
XX 1527
XX 1528
XX 1529
XX 1530
XX 1531
XX 1532
XX 1533
XX 1534
XX 1535
XX 1536
XX 1537
XX 1538
XX 1539
XX 1540
XX 1541
XX 1542
XX 1543
XX 1544
XX 1545
XX 1546
XX 1547
XX 1548
XX 1549
XX 1550
XX 1551
XX 1552
XX 1553
XX 1554
XX 1555
XX 1556
XX 1557
XX 1558
XX 1559
XX 1560
XX 1561
XX 1562
XX 1563
XX 1564
XX 1565
XX 1566
XX 1567
XX 1568
XX 1569
XX 1570
XX 1571
XX 1572
XX 1573
XX 1574
XX 1575
XX 1576
XX 1577
XX 1578
XX 1579
XX 1580
XX 1581
XX 1582
XX 1583
XX 1584
XX 1585
XX 1586
XX 1587
XX 1588
XX 1589
XX 1590
XX 1591
XX 1592
XX 1593
XX 1594
XX 1595
XX 1596
XX 1597
XX 1598
XX 1599
XX 1600
XX 1601
XX 1602
XX 1603
XX 1604
XX 1605
XX 1606
XX 1607
XX 1608
XX 1609
XX 1610
XX 1611
XX 1612
XX 1613
XX 1614
XX 1615
XX 1616
XX 1617
XX 1618
XX 1619
XX 1620
XX 1621
XX 1622
XX 1623
XX 1624
XX 1625
XX 1626
XX 1627
XX 1628
XX 1629
XX 1630
XX 1631
XX 1632
XX 1633
XX 1634
XX 1635
XX 1636
XX 1637
XX 1638
XX 1639
XX 1640
XX 1641
XX 1642
XX 1643
XX 1644
XX 1645
XX 1646
XX 1647
XX 1648
XX 1649
XX 1650
XX 1651
XX 1652
XX 1653
XX 1654
XX 1655
XX 1656
XX 1657
XX 1658
XX 1659
XX 1660
XX 1661
XX 1662
XX 1663
XX 1664
XX 1665
XX 1666
XX 1667
XX 1668
XX 1669
XX 1670
XX 1671
XX 1672
XX 1673
XX 1674
XX 1675
XX 1676
XX 1677
XX 1678
XX 1679
XX 1680
XX 1681
XX 1682
XX 1683
XX 1684
XX 1685
XX 1686
XX 1687
XX 1688
XX 1689
XX 1690
XX 1691
XX 1692
XX 1693
XX 1694
XX 1695
XX 1696
XX 1697
XX 1698
XX 1699
XX 1700
XX 1701
XX 1702
XX 1703
XX 1704
XX 1705
XX 1706
XX
```


Db 403 gctccgctcttgagtagcgggtctacgaltcaactcgtctactcgtctcagatgagccc 462
QY 1037 TATAACGACCCCTGGAGCAATACCTCGTTTATTAACGGATAGGGGCTCTTGGGCAAC 1096
Db 463 tataactacccttggagacaataaccctcgttataactcgtgataagggtactcttggcaac 522
QY 1097 ATCACTTGCATATGACCCCTGCAACAATGCGGAACGTCGCCAGCCGTGGAACCTCTAAC 1156
Db 523 atcaccttgcataacgcccctgcacaacatcgcgaaacgtcccaagccgaggaaaccttaac 582
QY 1157 AAGTGGAGCTCAACCGCGCAAGACTTCTTAATGCTAGAAATGCTACCCAGT 1216
Db 583 aagtcgagactcaaccgcccgaagacttccctaagtcgtctcgaaatcgactaccagaat 642
QY 1217 ATTCACTACTCAATGCGCATTTCAAGTATACCAACCGAATCAGACAAATAATATGAGCA 1276
Db 643 attcacctaccaatgcatcattcaagtaacacacacgaaatcagaacaaataatgagaca 702
QY 1277 AGTAGTAAAAAGATGCTGCTCTTACGCTGCTTTGCAANTGCTGGAACCTTACTATPAA 1336
Db 703 agtagtgaagaagatgctgctgtcttaccgcttgccttgcaaatggtggaaactactataa 762
QY 1337 CCATGTATATCCATTAAGTCCTTTAGTAGTGGAGTGAAGAGTCTCTAATGTC 1396
Db 763 ccaatgtatatacataaagtcgtctttagtgatgagtgagaaagatctctcaatgctc 822
QY 1397 GGAATGCTGTCATGGAAGAAAGACAGACCTATATGATGACCGACATGATGAAACAGTC 1456
Db 823 ggaactcgtgcataagaaagaaagacagcctataatgtagccgacatgatagaaacagtc 882
QY 1457 TTGACTTATGAACTGAGCAAAATGCTATCTGCTTGGCTCCCTCAGGCTGTAAACA 1516
Db 883 ttgacttatgaaactgagcaaaatgctatcttgccttgctccctcagctggttaaaaa 942
QY 1517 GGAACCTCTACATACAGACGGAATGGAATGAAACCAACATGACCTCTCAATTTGTA 1576
Db 943 ggaacctctaacatacagaagagaaatgaaatccacacatcgaagacttccatctgta 1002
QY 1577 GCACCTGATGACATATTTGCTGCTATACCGCTAATATTTCAATGCTGATGAGACAGC 1636
Db 1003 gcaacctgataactatctgtgtctatacgcgtataatcattcattgctatgtagacagc 1062
QY 1637 TATTTTAAACCGCTGACACCACTTTGATGCAATGCGCTTACGGCTGCCAAATTTAC 1696
Db 1063 tatictaaccgtctgcacacactttagtgcaatgagccttaagtcgtgtccaaagttaac 1122
QY 1697 CGCTCTATGATGACCTACCTGCTGGAAGGAAGCAATCCAAAGATTGGATATACCGAG 1756
Db 1123 cgtctatgtagctacccctgctgtaaggaagcaatccagaagattgaaatataccagag 1182
QY 1757 GGGCTCTACGAATGAGAAATCGTATTTAAATGATGCTCGTTTACGTGGA 1811
Db 1183 gggctctacagaatgagaaatcgtaattcaaaaaatggtgctgcttctacgtgga 1237

RESULT 9
AAZ35950
ID AAZ35950 standard; DNA: 1260 BP.
XX
AC AAZ35950;
DT 07-FEB-2000 (first entry)
XX
DE Streptococcus pneumoniae pbp1a TER isolate 1) nucleotide sequence.
XX
KW Streptococcus pneumoniae; penicillin binding protein: pbp2b; pbp1a;
KM transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
XX detection; identification; pneumococcal meningitis; ss.
OS Streptococcus pneumoniae.
XX
PN ZAG807024-A.

XX 28-APR-1999.
PD 05-AUG-1998; 982A-0007024.
XX
PF 01-AUG-1997; 972A-0006886.
XX
PR (SAME-) SOUTH AFRICAN INST MEDICAL RES.
XX (UYWI-) UNIV WITWATERSRAND.
XX (MEDI-) MEDICAL RES COUNCIL.
PI Klugman KP, Smith AM, Du Plessis M;
XX WPI: 1999-601770/51.
XX
DR Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis
XX
XX
PS Claim 11; Fig 4; 63pp; English.
XX
XX A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2b (pbp2b) gene
CC and the pbp1a gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC identified for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1a transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX
XX
SQ Sequence 1260 BP; 407 A; 293 C; 258 G; 302 T; 0 other;
Query Match 45.2%; Score 903; DB 20; Length 1260;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 575 GAACAGTATGAGAAGACGATCAATACCAATTAATGATGAGATCAATCAATCA 634
Db 1 gaacagtatgagaagacgacatcaacccaattactatgtagtaccacaagctccaatca 60
QY 635 GCAAGTATTAACCTGCTTACATGATGATTAATTAATTAATTAATTAATTAATTA 694
Db 61 gcaagtaattaccctgcttcaatgataattaccctcaaggaagatcaatcaatcaagttgaa 120
QY 695 GAAGAAACAGGCTTATACCTACTACACACTGGATGATGATGATGATGATGATGATG 754
Db 121 aagaagaacagcgtataaccgctcaacacggtgagatgttctacacaaatgtgagacaa 180
QY 755 GAAGCTCAAAAACATCTGCTGATATTTTACAAATACAGACGAATGCTGCTTACAGAC 814
Db 181 gaagctcaaaaacatctcgtggtgatatattcaatacagaagaatcgttgcctatccagac 240
QY 815 GATGAATTAAGTACGCTTACCATTTGATGATTTCTTAACGGTAAAGTATTAAGTCCAG 874
Db 241 gatgaattgcaagtcgcttctacacatgttgcattgattcctaagcgttaagtcattgccag 300
QY 875 CTAGAGACAGCCATCATGTAAGTAATGTTTCTTGGAAATTAACAGACATGAAGAAACA 934
Db 301 ctgagagacagcattcagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 360
QY 935 AACCGGACATGGGATTAATGAAGACGATCAACATGATGCTGCTGCTGCTGCTGCTG 994
Db 361 aaccggacatgggatatgaagacgatacgaacgatacgaacgatacgtcctcgttggagttac 420
QY 995 GGTGCTACGATTCAACTGCTACTATGTTACAGATGAGCCCTATTAATCACTCCCTGGACA 1054


```

Db 421 ggtgtctacgattcaactcgtactatcgttcaagatgagccctatactactccctggaca 480
OY 1055 AATACTCTCTGTTTATTAATGCGGATAGGGCTACTTTGGCAACATCATCTTGCANATACGCC 1114
Db 481 aatactccgttatatactacgtggaatagggctactcttgcaacatactccttgcaatagcgc 540
OY 1115 CTGCACCAATCGGAAAGCTCCAGCGGTGGAACCTTAACAGGTGCGACTCAACGCC 1174
Db 541 ctgcaacaacgcgaaacgtcccgccgtggaacactcaaaagctcgagcaacgcgc 600
OY 1175 GCCAAGACTTTCTTAATAGTGTAGAGATGACTACCAAGTATTCATCAATAGCC 1234
Db 601 gccaaagactctcctaatagtctcgaatcgcactaccgaatgtaactcactcaactgac 660
OY 1235 ATTTCAAGTAAACAAACCAATCAGACAAATAATATGAGCAAGTATGTAAGAAAGATGCT 1294
Db 661 attcaagaatacaacaacccgaatcagaacaaaataatgagcaagtagtgaagaatgagct 720
OY 1295 GGTGCTTACGCTGCTTTCGAATGCTGAGACTTACTATTAACCAATGTATATCCATAA 1354
Db 721 gctgtcttaagctcgtcttgcaaatggtgaaactactataaaccatgataatccataaa 780
OY 1355 GTGCTCTTATGATGAGGAGTGAAGAAAGAGTTCTTAATGTGCGAAGTGTGCGCATGAAG 1414
Db 781 gtcgtctttagtgaatgagtgaaagagctctcctaagctcgagcaactgctgcatgaag 840
OY 1415 GAAAGACAGCGCTATATGATGACCCAGATGATGAAAGAGCTTGTGACTTATGCACTGGA 1474
Db 841 gaaacaacagcctataatgataagcagacatgataaagacgtcttgatagtaagtaactgga 900
OY 1475 CCAAAATGCTATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1534
Db 901 cgaatgctactactctgctgctgctcctcagctggtgtaaaacggaaacctactataaa 960
OY 1535 GACGAGGAATTTGAAACCAACATCAAGACCTCTCAATTTGTAGCACCTGATGAATATTT 1594
Db 961 gacgaggaatgaaacacacatacagactctcaattgttagcacctgagcaactattt 1020
OY 1595 GCTGGCTATAGCGCTAAATATTCATGCTGATGACAGCGCTATTTCTTAACGCTGAGAA 1654
Db 1021 gctggcctaaacgcgtaaataatcattgctgataagcagagctatctcaccgctcgaca 1080
OY 1655 CCACCTGTATGGAATGCGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1714
Db 1081 ccacttgtaggaatgagccttaagctgctgctgcaaaagttaacgctctatgtagcctac 1140
OY 1715 CTGCTGGAAGAGCAATTCAGAGATTGGAATATPACAGAGGGGCTCTACAGAAATGGA 1774
Db 1141 ctgctcgaaggaagcaatcagagatggaatataccagaggggctctacagaaatgga 1200
OY 1775 GAATTCGATTTAAATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1834
Db 1201 gaatcgtatatttaaaaatggtgctgctcgtctcgtggaactcactgctccacaacacc 1260

```

RESULT 10
AA235951
ID AA235951 standard: DNA: 1260 BP.

AC AA235951;
DT 07-FEB-2000 (first entry)
XX Streptococcus pneumoniae pbp1A TER isolate m) nucleotide sequence.
DE Streptococcus pneumoniae: penicillin binding protein: pbp2B; pbp1A;
KW transpeptidase encoding region; TER: antibiotic resistance; diagnosis;
KW detection; identification; pneumococcal meningitis; ss.
OS Streptococcus pneumoniae.
XX
PN ZA9807024-A.

```

XX 28-APR-1999.
PD 05-AUG-1998; 982A-0007024.
XX 01-AUG-1997; 97ZA-0006886.
XX (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UTMI-) UNIV WITWATERSRAND.
PA (MEDI-) MEDICAL RES COUNCIL.
XX Klugman KP, Smith AM, Du Plessis M;
DR WPI: 1999-601770/51.
XX Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis
XX Claim 11; Fig 4; 63pp; English.
CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1A gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1A transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX Sequence 1260 BP: 407 A; 293 C; 258 G; 302 T; 0 other:
SQ

```

Query Match 45.2%; Score 903; DB 20; Length 1260;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

OY 575 GAACGATATGAAAGCACTCAATACCAATTTACTGATGAGGACTCAAGTCTCAATCA 634
Db 1 gaacgataagaaagcaagcaatcaacacatactgagactacaaagctcaatca 60
OY 635 GCAAGTATTTACCTGCTTACATGATGATATTTACTCAAGAGTCAATCAAGTTGAA 694
Db 61 gcaagtaattaccctgcttactatgataattacccaagaaagatcaatcaagttgaa 120
OY 695 GAAGAAACAGGCTATPACCTACTCAACACTGGAGATGATGTCTACCAAAATGTAGCCAA 754
Db 121 aagaagaacagctataacctgctcaaacctggatgaltgctacaaagtgaagccaa 180
OY 755 GAAGCTCAAAAACATCTGAGGATATTTTACATPACAGAGCAATAGCTTCCATCCAGAC 814
Db 181 gaagctcaaaaacatctgtggatatttacaatacagaagaaatgcttccatccaagc 240
OY 815 GATGAATTCAGAGTGCCTTCTACCAATTTGTTGATGTTTCTTAACGGTAAATCAATTC 874
Db 241 gatgaattgcagatgccttctacccattgttgatglttccaaaggttaagatcattgc 300
OY 875 CTAGAGACAGCCCATCAGTCAAGTAATGTTTCTCGAATTTAACCAACAGCAGTAGAAAC 934
Db 301 cttagagacagcccatcagcaagtaagtgttcccttcogagataacaaagcagtagaaaca 360
OY 935 AACCCGAGATGGGGATCACTATGAAACCGATCAGACAGATATGCTCTTGGATGATGAC 994
Db 361 aaccgcgactgggatacactatgaaacgcgatacagataatgctcctgcttgagttac 420
OY 995 GGTGCTACGATTCAGACTGCTACTATCGTTTCAGATGAGACCTATATACCTGAGACA 1054

```



```

Db      421  ggtgtcagcatcaactacgtactacgttcacgatgagccataactaaccctggagaca 480
OY      1055  AATATCTCTGTTTATATACGTGGAGGCTACTTTGGCAACATCACCTTGATACGCC 1114
Db      481  aatactcgtcttataacttgggataagggctacttggcaacacacacttgcataacgac 540
OY      1115  CTGCAACAATCGGAAACGTCGCCAGCCGTGAAACTTAAACAGTGGAGTCAACGCC 1174
Db      541  ctgcacaataatcggaacacgccccgacgctggaaacttaacaagcgcgagactaacgc 600
OY      1175  GCCAAGACTTTCTTAATGCTTGAAGATGACTACCCAGTATCTACTCAATGACC 1234
Db      601  gccaaagacttccctaaatggctcggatcgactacccaagtaactcaactcaatgacc 660
OY      1235  ATTTCAAGTAAACACCAAGCAATGACAAAAAATATGAGCAAGTACTGAAAAATGGCT 1294
Db      661  atttcaagtaacacacacgacgatacgaacaaaataataggaagtagtgaagaatgact 720
OY      1295  GCTGCTTACGCTGCTTGGCAAAATGCTGAACCTTACTATTAACCAATGATATCAATAA 1354
Db      721  gctgcttaacgctgcttctgtaaaatgctggaacttactataacacatgatatctcataa 780
OY      1355  GTGCTCTTTTGTGATGGAGTGAAGAAAGAGTTCTTAATGTCGGAAGTCTGCTCATGAG 1414
Db      781  gtcgtcttctagtagtggagtagaagaagttctcttaatgtcggaaactcgtgcacatga 840
OY      1415  GAAACGACAGCTTATGATGAGACGACATGATGAAAAACGCTTACTGATTTGGAACGGA 1474
Db      841  gaaacaaacgctataatgatacgacacatgataagaacagctcgtgagttatggaactgga 900
OY      1475  CGAAATGCTATCTTGGCTTGGCTCCCTCAGGCTGGTAAACAGGAACCTCTACTATACA 1534
Db      901  cgaatgctcctactctgttctgctccctcaggtcgtgtaaaacaggaacctcttaactaca 960
OY      1535  GACGAGGAATTTGAAAAACCATCAACACCTCTCAATTTGTAGACCTGATGAACATTTT 1594
Db      961  gacgaagaagaattgaaacacacataagaacctctcaattttagaaccttgcgaacttct 1020
OY      1595  GCTGCTATACGCGTAAATTTTCAATGGCTGATGAGACGAGCTTTTCAACCCGCTGACA 1654
Db      1021  gctggtcattcgcgttaataatctcaatgctgctgataagcaagctcttctaacgctcgaca 1080
OY      1655  CGACTTGTAGCAATGCGCTTACGCTGCGTCCAAAGTTTACCGCTCTATGATGACCTAC 1714
Db      1081  ccaacttgtagcaatgagccttaacgctcgcgcgaagtttaacgctctatgatgacctac 1140
OY      1715  CTGCTTGAAGAGCAATCCAGAAAGATTGGAATATACCAAGAGGGGCTCTACAGAAATGGA 1774
Db      1141  ctgcttgaagaagaacatccagagattggaatactaccagagggcctctacagaatgga 1200
OY      1775  GAATTCGTAATTTAAAAATGCTGCTGCTTACGTCGTAACGTAACCTGCTCAACAACAACC 1834
Db      1201  gaattcgattataaaaatgctgctcgttctacgttgaactcaccctgctccacaacaacc 1260

```

RESULT 11
AAZ35939
ID AAZ35939 standard; DNA; 1260 BP.

XX AAZ35939;
XX
XX 07-FEB-2000 (first entry)
DE Streptococcus pneumoniae pbp1A TER isolate a) nucleotide sequence.
XX
XX Streptococcus pneumoniae; penicillin binding protein; pbp2b; pbp1A;
KW transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
XX detection; identification; pneumococcal meningitis; ss.
OS Streptococcus pneumoniae.
XX
XX Location/Qualifiers
FH key

```

FT      CDS      1..1260
FT      FT      /*tag= a
FT      FT      /product= "pbp1A TER"
FT      FT      /note= "no stop codon given"
FT      FT      ZAN9807024-A.
XX      XX
XX      XX      28-APR-1999.
XX      XX
XX      XX      05-AUG-1998; 98ZA-0007024.
XX      XX
XX      XX      01-AUG-1997; 97ZA-0006886.
XX      XX
XX      XX      (SAME-) SOUTH AFRICAN INST MEDICAL RES.
XX      XX      (UYWI-) UNIV WITWATERSRAND.
XX      XX      (MEDT-) MEDICAL RES COUNCIL.
XX      XX
XX      XX      Klugman KP, Smith AM, Du Plessis M;
XX      XX      WPI; 1999-601770/51.
XX      XX      P-PSDB; AAY56106.
XX      XX
XX      XX      Polymerase chain reaction assays for detecting Streptococcus pneumonia
XX      XX      useful for the diagnosis of pneumococcal meningitis
XX      XX      Claim 11, Fig 4; 63pp; English.
XX      XX
XX      XX      A polymerase chain reaction (PCR) assays have been developed for
XX      XX      detecting an antibiotic resistant strain of Streptococcus pneumoniae
XX      XX      using primers based on the penicillin binding protein 2b (pbp2b) gene
XX      XX      and the pbp1A gene. The products and methods can be used for detecting
XX      XX      S. pneumoniae, particularly antibiotic-resistant strains. They can be
XX      XX      used for simultaneously diagnosing pneumococcal meningitis and
XX      XX      identifying any antibiotic-resistant S. pneumoniae strains in a sample.
XX      XX      The methods can be used for detecting S. pneumoniae strains resistant
XX      XX      to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
XX      XX      The assays can be adapted to detect other pathogens causing meningitis.
XX      XX      The assays can be used to detect an antibiotic resistant strain of
XX      XX      S. pneumoniae with a minimum inhibitory concentration (MIC) of
XX      XX      0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
XX      XX      a 224 bp product. The present sequence represents a Streptococcus
XX      XX      pneumoniae pbp1A transpeptidase encoding region (TER) isolate nucleotide
XX      XX      sequence from the present invention.
XX      XX
XX      XX      Sequence 1260 BP; 405 A; 294 C; 261 G; 300 T; 0 other;
XX      XX
XX      XX      Query Match      41.9%; Score 838; DB 20; Length 1260;
XX      XX      Best Local Similarity 99.4%; Pred. No. 0;
XX      XX      Matches 1188; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX      XX
OY      617  CTACAAAGTCTCAANTAGCAAGTAATTTACCTGCTTACATGATGATTAATCTCAAGAA 676
Db      43  ctacaagcttcaaatcagcaagaatlaacccgtcttaactgataatlaaccctcaagaa 102
OY      677  GTCAATCAATCAAGTTGAAGAAAGAAAGCGCTATTAACCTACTCACACTGGAGATGTC 736
Db      103  gtcatcaatcaagttgaagaagaacagcgtcataactactcaacaactgagatgacgctc 162
OY      737  TACCAATATGTAGACCAAGAGCTCAAAACATCTGTGGGATATTTTCATACAGAGAA 796
Db      163  tacacaatgtlagaccagaagctcaaaaacatctggtggatattacaatacagaagaa 222
OY      797  TACGTCCTATCCAGACGATGAATTTCAAGTGCCTTCTACATGTTGATGTTTCAAC 856
Db      223  taagtgcctatccagacgagatgaaatgcaagtgctctaccatgttgatggttctcaac 282
OY      857  GGTAAAGTCAATTTGCCAGCTAGAGCAAGCCATCAGTCAAGTAATGTTCTTGGAAAT 916
Db      283  ggtaaagtcaattgcccagctagagagcagcatcagtaagtaatgttctcctcgatc 342
OY      917  AACCAAGCAGTAGAACAACCGGAGCTGGGATCAACTGTGAACCGATCACAGACTAT 976

```



```

Db 343 aaccaaagcagtagaanaaaacccgcgactggygagtaacatagaacagatcaacagactat 402
Oy 977 GCTCGTGGCTGGAGTAGCGGTGTCTACGATTCAGTACTGCTTACGTTACGATGAGCCG 1036
Db 403 gctccgcctctggagtagcgtgtctacgagtaacactgcactatcgttcaagatgagccc 462
Oy 1037 TATTAACCTACCCGGAGCAAAATACCTCCTGTTTATTAACCTGGGATAGGGGCTACTTTGGCAAC 1096
Db 463 tataactaccctggagaataataccctgtttataactgagataggggtcactttgccaac 522
Oy 1097 ATCACTTTCGATATACCCCTGCGACAAATGCGGAACAGTCCGACCGCTGGAAACTCTAAC 1156
Db 523 atcaactctgacaacacgcttgacaacatcgcaaaacgtccagccgttggaaactcttaac 582
Oy 1157 AAGGTGAGACTCAACCGCCCAAGACTTTCCTAAATGCTGTGAGATTCAGTACCAAGT 1216
Db 583 aaggtcgagactcaaccgcgcagagacttccctaataatgctcggaaactgactacccaagt 642
Oy 1217 ATTCACTACTCAAAATGCCATTTCCAACTAACACAAACCGAATCAGACAAAATAATGAGAGA 1276
Db 643 attcaactcaaatgacatcattcaagtaacacacacgaaatcagacaanaataatgagaca 702
Oy 1277 AGTAGGAAAAAGATGGCTCTGCTTACGCTGCTTGGCAATGCTGGAACCTACTATAA 1336
Db 703 agtagtaaaaagatggtctgtctacgcttgacgtcttgcaaatggttgaacttactataa 762
Oy 1337 CCAATGTATTCATTAAGTTCGTTTACTGATGGAGGAAAAAGATGTTCTATATGTC 1396
Db 763 ccaagtataatccataaagtcgctcttagtagaggaagaaagatccatccatagtc 822
Oy 1397 GGAATCTGTGCGAGGAAAGAACGACAGCGCTATATGATGACGACATGATGAAAAACAGTC 1456
Db 823 ggaactcgtgcacatgagaagaacagacagcctatatgatgacgcacatgataaagcagtc 882
Oy 1457 TTAGCTATGAGAACGAGGAAATGCTATGCTTGTGCTGCTCCCTGAGCTGTTAAAGA 1516
Db 883 ttgagtagtaagaaacgagcaaaatgctactctgtctgctccctcaaggcttgtaaaaca 942
Oy 1517 GGAACCTCTAATATACAGACGAGAAATGGAATGAAACACATCAAGACCTCTCAATTTGTA 1576
Db 943 ggaacctcaactataacagagaggaatgtgaaacacacatacgaagcctcaattgtga 1002
Oy 1577 GCACCTGATGAACATTTTCTGCTGCTATACGCGTAATATTCATATGCTGTATGAGACAGC 1636
Db 1003 gaacctgagaaactatltgtgctatacgcgttaaatatcaatgactgctatgagacaagc 1062
Oy 1637 TATTCTAACCGTCTGACACCACTTGTAGGCATAGCGCTTACCGTCCCTGCCAAAGTTTAC 1696
Db 1063 tatctaaacgcttgacacactttaggcaatgaccttaccgctgcgtccaaagtttac 1122
Oy 1697 CCCTATGATGACCTACCTGCTGTGAAGAGCAATCCAGAGATTGGAATATACCAAG 1756
Db 1123 cgtctcatgatagaactactgtctgtaagaagaatccagaagatctggaatataccagag 1182
Oy 1757 GGCGCTTCAGAAATGAGAAATTCGATTATTAATAATGATGCTGTTCTACGTGGA 1811
Db 1183 gggctctacagaatgaggaattcgtatttaaaaatcgtgcctcgttctacgtgga 1237

```

RESULT 12
AAZ35947
ID AAZ35947 standard; DNA; 1260 BP.

AAZ35947;

07-FEB-2000 (first entry)

Streptococcus pneumoniae pbp1A TEF isolate 1) nucleotide sequence.

Streptococcus pneumoniae: penicillin binding protein; pbp2B; pbp1A; transpeptidase encoding region; TEF; antibiotic resistance; diagnosis; detection; identification; pneumococcal meningitis; ss.

```

OS Streptococcus pneumoniae.
XX
PN ZA9807024-A.
XX
PD 28-APR-1999.
XX
PF 05-AUG-1998; 982A-0007024.
XX
PR 01-AUG-1997; 972A-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UWI-) UNIV WITWATERSRAND.
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
DR WPI; 1999-601770/51.
XX
PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis
XX
PS Claim 11; Fig 4; 63pp; English.
XX
CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1A gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1A transpeptidase encoding region (TEF) isolate nucleotide
CC sequence from the present invention.
XX
SQ Sequence 1260 BP; 408 A; 286 C; 264 G; 302 T; 0 other;

```

Query Match 41.2%; Score 824; DB 20; Length 1260;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 824; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 617 CTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGAGATTAATTAAGTCAAGAA 676
Db 43 ctacaagctcacaatcagaagaatlaattaccctgtctacatgataatlaaccctcaagaa 102
Oy 677 GTCATTAATCAAGTTGAACAGAAACAGGCTAATACCTCTACACTGCGATGGAGATGTC 736
Db 103 gtcatcaatcaagttgaaagaagaacaggtctaaactctcaacgtggaagatgctc 162
Oy 737 TACACAAATGTAGACCAAAAGCTCAAAAACATCTGTGAGATTTATACATACAGCGAA 796
Db 163 tacacaatgtagaccacaagaagctcaaaaacatctgtggaatatctaacataacagcgaa 222
Oy 797 TACGTTGCCTATTCAGACGATGAATTCGAAGTGCCTTACCAATTTGATGTTTAAAC 856
Db 223 tacgttgcctatccagaagatgaattgcaagtgcgtcttaccatctgtgatgtctaac 282
Oy 857 GGTAAGTCTATGCTCCAGCTAGAGACAGCCCATCATCAAGTAATTTCTCCGGAAT 916
Db 283 ggtlaagttatgtccagctagagcagccatcagtaagaatgttccctcggaaat 342
Oy 917 AACCAAGCATAGAAACCAACCGCAGCTGGGATGCAACATGATGAACCGATCACAGACTAT 976
Db 343 aaccaaagcagtagaanaaaacccgcgactggygagtaacatagaacagatcaacagactat 402
Oy 977 GCTCGTGGCTGGAGTAGCGGTGTCTACGATTCAGTACTGCTTACGTTACGATGAGCCG 1036

```


Db 403 gccctgccttgagtagcaggtgtctacagatccactacgtctactacgtacagatgagccc 462
QY 1037 TATACTACCCCTGGGGAACAAATACCTCGTTTATACCGGGATGGGGCTACTTGGCAAC 1096
Db 463 tataactaaccccyggaacaatactccgtttataaactgggagtaggggctactttggcaac 522
QY 1097 ATCACTTGGCAATACCCCTGCAACAATCGCAAGCTCCCAACCCGTGGAAGCTCTAAAC 1156
Db 523 atcaacttgataatagccctcgcaacaactcgcaagaacgtccacgacgtggaactctaaac 582
QY 1157 AAGGTGGACTCAACCGCGCAAGACTTTCTTAATGCTTAGAATGCTACTGCCAAGT 1216
Db 583 aaggtcggaactcaaccgcgcgaagacttctctaattgtctaaagatcgactaccacaagt 642
QY 1217 ATTACATACCAATGATGCAATTCATTAAGTAACACACCGAATCACAAATAATATGAGCA 1276
Db 643 atcaactactcaaatgcattcaatgaataacaacgcgaatcgaacaataatggaagca 702
QY 1277 AGTAGTGAAGAAGATGGCTGCTTACGCTTTCGCTTGGCAATGGTGGAACTTACTATAA 1336
Db 703 agtagtgaagaagatgctgtctgtctacgcttgcaatggtggaacttactataaa 762
QY 1337 CCATGTATATCCATAAAGTCGCTTTAGTAGAGGAGTGAAAAAGATTCTTAATGTC 1396
Db 763 ccaatgtataatccataaagtcgtcttaagtatgagtagtgaagaaggtctctaaatgtc 822
QY 1397 GGAAGTCGTGCATGAAGAAAGACAGCCTATATGATGACGCA 1440
Db 823 ggaactcgtgcatagaagaagaacagcctataatgatgacga 866

RESULT 13
AA083241
ID AA083241 standard; DNA; 960 BP.

XX AA083241;

XX 14-FEB-1996 (first entry)

DE Streptococcus pneumoniae strain SPRU42 Exp2 DNA.

KW Exp2: export protein; permease like protein;
KW penicillin binding protein 1A; pbp1a; pbp1; exp1; exp3; pad1;
KW pyruvate oxidase; virulence determinant; regulatory element;
KW acellular vaccine; antibody; ds.

OS Streptococcus pneumoniae.

PN W09506732-A2.

PD 09-MAR-1995.

PF 01-SEP-1994; 94WO-US09942.

PR 01-SEP-1993; 93US-0116541.

PA 18-MAY-1994; 94US-0245511.

PA (UYRO) UNIV ROCKEFELLER.

PI Masure HR, Pearce BJ, Tuomanen E;

DR WPI; 1995-115448/15.

DR P-PSDB; AAR70153.

PT Novel gene fragments encoding specific bacterial exported proteins

PS - specifically of S. pneumoniae; useful as vaccines

XX Claim 28; Page 88-9; 168pp; English.

CC This sequence represents the DNA encoding exp2. This sequence is
CC identical to that for ponk which encodes penicillin-binding protein 1A
CC (pbp1a). The protein encoded by this sequence is involved in adhesion
CC of bacteria to target cells. This sequence encodes an exported protein

CC of S.pneumoniae. Export proteins are the proteins in pathogenic
CC bacteria that are virulence determinants. Other export proteins include
CC pbp1a (see AAR70152), exp1, exp3, and pad1 (encoded by the sequence shown
CC in AA083259). This sequence can be inserted into an expression vector
CC (preferably a bacterial expression vector) to provide for high levels of
CC expression of the protein. The protein can then be used in the
CC production of an acellular vaccine. These vaccines are used to provide
CC protection from Gram positive bacterial infection. Antibodies against
CC export proteins can be used for diagnosis of infection and in passive
CC immune therapy.

XX Sequence 960 BP; 315 A; 224 C; 182 G; 239 T; 0 other;

Query Match 39.8%; Score 796; DB 16; Length 960;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAAATTAATCACTCAATGCTGACTTGGTTCGAAACCGCGCTCAA 60
Db 12 taaatctacgacaataaataatcaactcattgctgacttgggttcgaacgcgcgtcaaa 71
QY 61 TGCCCAAGCTAATGATATTTCCACAGATTGGTTAAGCAATGCTTTCTATCGAAGCA 120
Db 72 tgcaccaagctaaatgatatccacagatttggtaaaggcaatggttctatcgaagacca 131
QY 121 TCGCTTCTTGACACACAGGGGGATTTGATACATCCGATCCGTGGAGCTTCTTGGCAA 180
Db 132 tgccttcttcgacacaggggagattgataccatccglatatccgttggagcttcttgcgcaa 191
QY 181 TGTGCAAAAGCAATTCCTCCAGGTGATCACTCTACCCCAAGTTGATTAAGTTGAC 240
Db 192 tctgcaaaagcaattccctccaagttgattcagctctcaacaaagttgattgaattgac 251
QY 241 TTACTTTTCACTTCGACCTTCGACCAAGTATTTTCGTAAGGCTCAGGAAGCTTGGT 300
Db 252 ttactttcaactcgactccgacagactatttccgtaaaggctaaaggacttggct 311
QY 301 AGCGATCACTTGAACCAAAACCAACCAAGCAAGAAATCTTGACCTACTATTAATAA 360
Db 312 agcgattcagttagaacaaagaacaaacaaagaagaattctgacctacataataaa 371
QY 361 GGTCTACATGCTGAATGCAAGTATGGAATGCAAGACAGCTCAAAATCTATGATGTA 420
Db 372 ggtctacatgcttaatggaactatggaactatggaactatggaactatggaactatgga 431
QY 421 AGACCTCAATATTTAAGTTTACCTCACTTACCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 432 agacctcaataatlaaagtttacctcagttagcttctgctgctgctgctgctgctgctg 491
QY 481 AAACCAATATGACCCCTATTCACATCCAGAGGACGCGCAAGCGGAAACTTGGCTT 540
Db 492 aaaccaaataagaccctatttcaacatccagaagaagcccaagacgcgcgaacttggctt 551
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAACAGTCAAT 600
Db 552 atctgaaatgaaaaatcaaggtctacatctctgctgcaagctatgagaagcagctcaaac 611
QY 601 ACCAATTTAGTATGAGTCAAAAGTCTCAATATGCAAGTCAATTTACCTCTTACATGA 660
Db 612 accaatttactgattggtctacaaggtctcaaatcaagaagtaattaccctgcttaccag 671
QY 661 TAAATACCTCAAGAGTCAATCAATCAAGTTGAAGAAGAAAGAGCTATTAACCTACTCAC 720
Db 672 taattacctaaagaagatcctacatcaagttgagaagaagaacgctactactactcac 731
QY 721 AACTGGATGATGCTTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
Db 732 aactggatgatgcttaccacaatgtagaccaagagctcaaaaacatctgtyggatatt 791
QY 781 TTACATACAGAGATAGTGTGCTTATCCAGACGATGAATTTGCAAGTCCGCTTACAT 840
Db 792 ttacaatacagacgaatagcttgcctatccagacgataatgtaattgcaagcttctacac 851

QY 841 TGTGATGTTTCTAACGTTAAAGTCATTTGCCAGCTAGAGACGCCATCAGTCAAGTAA 900
|||||
DB 852 tgttgatgtttcttaacgagtaagtcattgcccagctagagacgcgcacatcagtaagtaa 911
|||||
QY 901 TGTTCCTTCGGAATTAAACCAAGCAGTAGAACAACCCGCGACTGGGGA 949
|||||
DB 912 tgttccttcgagattaccagcagtagaacaacccgactgggga 960
|||||
RESULT 14
AAZ35948
ID AAZ35948 standard; DNA; 1260 BP.
AC AAZ35948;
XX
XX
DT 07-FEB-2000 (first entry)
XX
XX Streptococcus pneumoniae pbp1A TER isolate j) nucleotide sequence.
XX
KM Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;
KM transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
KM detection; identification; pneumococcal meningitis; ss.
XX
OS Streptococcus pneumoniae.
XX
PN ZA9807024-A.
XX
XX 28-APR-1999.
XX
XX 05-AUG-1998; 982A-0007024.
XX
PR 01-AUG-1997; 972A-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UWI-) UNIT WITWATERSRAND.
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
XX
XX WPI; 1999-601770/51.
XX
XX Polymerase chain reaction assays for detecting Streptococcus pneumonia
XX useful for the diagnosis of pneumococcal meningitis
XX
PS Claim 11: Fig 4; 63pp: English.
XX
XX A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1A gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1A transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX
XX Sequence 1260 BP; 403 A; 285 C; 264 G; 308 T; 0 other;

Query Match 38.2%; Score 764; DB 20; Length 1260;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 864; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 575 GAACAGTATGAGAAAGCAGCTCAATACCAATTAATGATGAGACTACAAAGTCTCAATCA 634
|||||

DB 1 gaacaglatgagaaagcagctcaatcaaccataactgactgagctacacaaagctcacaatca 60
QY 635 GCAGATTAATTACCTGCTTACATGATGATTAATTACCTGAGAGATCAATCAAGTTGAA 694
|||||
DB 61 gcaagtaattaccctgcttacaatgataatcaatcaagaagatcaatcaatgaa 120
|||||
QY 695 GAAGAAACAGGCTTAATACCTACTACACACTGGAGATGATGTCACAAATGTAGACCA 754
|||||
DB 121 gaagaaacagcgtataaccctgctcacaactggagtgatgctcacaatgtagccaa 180
|||||
QY 755 GAAGCTCAAAAACATCTGTGGGATATTTCATATACAGACGATAGCTTGCTTCCAGAC 814
|||||
DB 181 gaagctcaaaaacatctgtggtatcttaacaacagcaatagctgtgctatccagac 240
|||||
QY 815 GATGAATTCGAATGCTGCTTACCATTTGATGATTTGATGATTAAGCGTAAAGCATGGCCG 874
|||||
DB 241 gatgaattcgaaatgcttcttaccattgttgaatgcttcaagctaaagctatggccg 300
|||||
QY 875 CTAGAGACAGCCATCAGTCAAGTAAATGTTCCCTTCGGAATTAAACAGCAGTAGAACA 934
|||||
DB 301 ctagagacagccatcagctcaatgataatgcttcttgaatcaacaagagtagaagaa 360
|||||
QY 935 AACCGGACTGGGATCAACTATGAACCGATACAGACTATGCTGCTTGGAGTAC 994
|||||
DB 361 aaccggactgggagatcaactatgaacacgatacagactatgcttcttggagtagc 420
|||||
QY 995 GGTGCTAGATCACTGCTATGCTATGCTACAGATGAGCGCTTAACTACCTGGGACA 1054
|||||
DB 421 ggtgctagatcactgcttactacgcttcaagatgaagccataactacccctgggaca 480
|||||
QY 1055 AATACTCTGTTTAACTAGTGGATAGGAGGCTACTTTGGCAACATCACTTGCAATACGCC 1114
|||||
DB 481 aatactctgttttaactatgagtagggctacttggcaacatcaacttgaatagcgc 540
|||||
QY 1115 CTGCAACATCGCGAAGACGTCACCGCGTGGAACACTTAACAAGTGGAGCTCAACGCC 1174
|||||
DB 541 ctgcaacatcgcgaaacgctccagccgtggaacacttaacaaagctggagctcaacgcg 600
|||||
QY 1175 GCCAAGACTTCTTAATGCTAGGAATGAGATGATACCAAGTTCATCACTCAATAGCC 1234
|||||
DB 601 gccaaagacttcttctaaatgctcggatcgaatcgactaccgaatctcaactcaaatgccc 660
|||||
QY 1235 AATTCAAGTAAACACACCGAATCAGACAAAATATGAGCAAGTAGTGAAAGATGCT 1294
|||||
DB 661 attcaagtaacacacacgcaatcagacaaaataatgagcaagtagtgaagaagagct 720
|||||
QY 1295 GCTGCTTACGCTGCTTTCGAAATGGTGAACCTTACTATTAACCAATGTATATCAATAA 1354
|||||
DB 721 gctgcttacgctgcttgcataatggtggaacttactataaccaaagtatatacatcaa 780
|||||
QY 1355 GTGCTTTTATGATGAGAGTGAAGAAAGATTCTTAATGTGGAACCTGTCGCTATGAAG 1414
|||||
DB 781 gtgctttttagatgagtaggagtagaaagattcttaactcgaactcgtgctcagaag 840
|||||
QY 1415 GAAGACAGACCCATATGATGATGACCA 1440
|||||
DB 841 gaagacagacccatataatgatgaccca 866
|||||
RESULT 15
AAZ35949
ID AAZ35949 standard; DNA; 1260 BP.
XX
XX AAZ35949;
XX
XX
DT 07-FEB-2000 (first entry)
XX
XX Streptococcus pneumoniae pbp1A TER isolate k) nucleotide sequence.
XX
XX Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;
XX transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
XX detection; identification; pneumococcal meningitis; ss.
XX

OS Streptococcus pneumoniae.
 XX
 PN ZA9807024-A.
 XX
 PD 28-APR-1999.
 XX
 PF 05-AUG-1998; 982A-0007024.
 XX
 PR 01-AUG-1997; 972A-0006886.
 XX
 PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
 PA (UYWI-) UNITY WITWATERSRAND.
 PA (MEDI-) MEDICAL RES COUNCIL.
 XX
 PI Klugman KP, Smith AM, Du Plessis M;
 DR WPI; 1999-601770/51.
 XX
 PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
 useful for the diagnosis of pneumococcal meningitis
 PS
 PS Claim 11, Fig 4; 63pp; English.
 XX
 CC A polymerase chain reaction (PCR) assays have been developed for
 CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
 CC using primers based on the penicillin binding protein 2b (pbp2b) gene
 CC and the pbp1a gene. The products and methods can be used for detecting
 CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
 CC used for simultaneously diagnosing pneumococcal meningitis and
 CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
 CC The methods can be used for detecting S. pneumoniae strains resistant
 CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
 CC The assays can be adapted to detect other pathogens causing meningitis.
 CC The assays can be used to detect an antibiotic resistant strain of
 CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
 CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
 CC a 224 bp product. The present sequence represents a Streptococcus
 CC pneumoniae pbp1a transposase encoding region (TER) isolate nucleotide
 CC sequence from the present invention.
 CC
 XX
 SO Sequence 1260 BP; 411 A; 282 C; 260 G; 307 T; 0 other;

Query Match 37.3%; Score 745; DB 20; Length 1260;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 845; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 594 TCATATACCAATCTAGTATGATGATATCAAGTCTCAATGACGATATATACCTGCTT 653
 DB 20 tcaatcaccaatctagatgactacaagaatcacaatcagaagaatcaatcctgctt 79
 QY 654 ACATGATATTAATCTCAAGAGATCATCAATCAAGTTGAAGAAGAAAGCGTATATAC 713
 DB 80 acatgataatctacccaagaagatcacaatcagaatgagaagaagaacgctataacc 139
 QY 714 TACTCACAACCTGGATGATGCTTACACAATGTAGACCAAGAGCTCAAAAACATCTGT 773
 DB 140 tgcacaacaactggatgactgctacacaatgtagacccaagaagctcaaaaacatctgt 199
 QY 774 GGGATATTTACATACGAGAGATAGCTTGCCTATCCAGAGATGAATTCGAAGTCCGTT 833
 DB 200 gggatatttacaatacagaagaatagcttgcctatccagacgaatgaaatgcgctt 259
 QY 834 CTACGATTTGATGATGCTTCAACGTAAGTCAATGCGCCGCTGAGACACCGCATACAGT 893
 DB 260 ctacgatctgtgactgcttcaacgtaaaagtcacgctcagtagagacgcacatcagt 319
 QY 894 CAAGTATATGTTCTTCTGGAATTAACCAACAGTAGAAGAAACCGGAGCTGGGATCAA 953
 DB 320 caagtataatgttcttcttggaattaaccaacagtagaagaaacacgagctgggatacaa 379
 QY 954 CTATGAAACCGATACAGACTATGCTCTGCTTGGAGTAGGATGCTGTACGATTTCAACTG 1013
 DB 380 ctatgaaacccgatcacagactatgctctgcttggagtagaagtgctacatccaactg 439
 QY 1014 CTACTATCGTTCACGATGAGCCCTATTAATCTACCTGGAGCAAAATACCTGTTATACT 1073
 DB 440 ctactatcgttccacgaagagccctatacctcctggagcaaaatacctcgtttataact 499
 QY 1074 GGGATAGGGGCTACTTTGGCAACATCACTTGAATAGCCCTGCAACATGCGGAAACG 1133
 DB 500 gggataggggctacttttggcaaacataccttgaataacgaccttgcaacaatcgcgaaacg 559
 QY 1134 TCCCAAGCCGTGGAACCTCTAAACAAGTGGAGCTCAACCGGCCCAAGACTTCTTAATG 1193
 DB 560 tcccaagccgttgaaaccttaacaaggctcggaatccacgcgcgcaagaacttcttaaatg 619
 QY 1194 GTCTAGAGATGAGTACCCAGATATTCATCTACTCAATGACCAATTTCAAGTAAACAACG 1253
 DB 620 gtctagagatgagtagtaccacgaagtattcactactcaaatgcatcctaagtaacacacg 679
 QY 1254 AATGAGACAAAAAATATGAGACCAAGTAGTGAAGAAAGATGCTGCTTACGCTCCCTTG 1313
 DB 680 aatgagacaaaaaatatgagacccaagtagtgaagaagatgctgctgcttaagctgcttgg 739
 QY 1314 CAATGTGTGCACTTACTATTAACCAATGTTATTCATTAAGTCTTTAGTATGGA 1373
 DB 740 caatgtgtgcacttactatataaccaaagtatataccaataagctgctttagtggga 799
 QY 1374 GTGAAAAAGAGTTCTTAATGTGGAACGTGTCATGAAGAGAAACGACAGCCTATATGA 1433
 DB 800 gtgaaaaagagttcttctaagtgcgaactcgtgcacaggaagaacgacagccatatga 859
 QY 1434 TGACCGA 1440
 DB 860 tgaccga 866

RESULT 16
 AAH02060
 ID AAH02060 standard; DNA; 930 BP.
 XX
 AC AAH02060;
 XX
 DT 24-JUL-2001 (first entry)
 XX
 DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2053.
 XX
 KW Species specific; genus specific; family specific; probe; detection;
 KW identification; algal; archaeal; bacterial; fungal; parasitical;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 KW vaccine; primer; ds.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200123604-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-CA01150.
 XX
 PR 28-SEP-1999; 99CA-2283458.
 PR 19-MAY-2000; 2000CA-2307010.
 XX
 PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
 XX
 PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;
 XX
 DR WPI; 2001-245006/25.
 XX
 PT Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 PT archaeal, bacterial, fungal and parasitcal species in a test sample -

XX Disclosure; Page 1468; 1580bp; English.
XX
PS
CC The present invention describes a method for generating a repository of
CC nucleic acids of tuf, fus, atp and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitical
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archaeal, bacterial, fungal
CC and parasitical species, genus, family and group. A nucleic acid (I)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexa nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (I) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp. using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
XX
SQ Sequence 930 BP; 289 A; 212 C; 196 G; 233 T; 0 other;

Query Match 31.0%; Score 620; DB 22; Length 930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 670; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 770 CTGTGGGATATTTCACATACAGACGAATACGTTGCTATCCAGCATGATGATTCAGATC 829
DB 1 CTTGGGATATTTCACATACAGACGAATACGTTGCTATCCAGCATGATGATTCAGATC 829
OY 830 GCTTACCATGTTGATGATGTTCTACAGGTAAGATGATGATGATGATGATGATGATGAT 889
DB 61 GCTTACCATGTTGATGATGTTCTACAGGTAAGATGATGATGATGATGATGATGATGAT 889
OY 890 CAGTCAAGTAAATGTTCTTCGGAATTAACCAAGAGTAAGAAACCAACCGCATGGGGA 949
DB 121 CAGTCAAGTAAATGTTCTTCGGAATTAACCAAGAGTAAGAAACCAACCGCATGGGGA 949
OY 950 TCACATATGAACCGATTCACAGACTATGCTGCTTGAGTACGATGATGATGATGATGAT 1009
DB 181 TCACATATGAACCGATTCACAGACTATGCTGCTTGAGTACGATGATGATGATGATGAT 1009
OY 1010 ACTGCTACTATGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1069
DB 241 ACTGCTACTATGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1069
OY 1070 AACTGGGATAGGGGCTACTTTGGCAACATACCTTGGCAATACGCTCGGCAACATGCGA 1129
DB 301 AACTGGGATAGGGGCTACTTTGGCAACATACCTTGGCAATACGCTCGGCAACATGCGA 1129
OY 1130 AACGTCGCCCGCTGGAACCTGAACCAAGTCGACCTAACCGCGCCCAAGACTTTCCTA 1189
DB 361 AACGTCGCCCGCTGGAACCTGAACCAAGTCGACCTAACCGCGCCCAAGACTTTCCTA 1189
OY 1190 AATGCTCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1249
DB 421 AATGCTCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1249
OY 1250 ACCGATATGAGCAAAAATATGAGCAAGTATGTAAGAAAGATGGCTGCTTACGCTGCC 1309
DB 481 ACCGATATGAGCAAAAATATGAGCAAGTATGTAAGAAAGATGGCTGCTTACGCTGCC 1309
OY 1310 TTTCGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1369

DB 541 TTTCGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
OY 1370 GGGAGTGAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1429
DB 601 GGGAGTGAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
OY 1430 ATGATGACCGA 1440
DB 661 ATGATGACCGA 671

RESULT 17
AAH01023
ID AAH01023 standard; DNA; 1199 BP.
XX
XX AAH01023;
AC
XX 24-JUL-2001 (first entry)
XX
XX Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1014.
DE
XX
XX Species specific; genus specific; family specific; probe; detection;
KW identification; algal; archaeal; bacterial; fungal; parasitical;
KW microorganism; diagnosis; translation elongation factor Tuf; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
XX Streptococcus pneumoniae.
OS
XX WO200123604-A2.
PN
XX 05-APR-2001.
PD
XX 28-SEP-2000; 2000WO-CA01150.
PF
XX 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
PA
XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
XX WPI: 2001-245006/25.
PS
XX
XX Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitical species in a test sample -
XX
XX
XX Claim 27: Page 968; 1580bp; English.
PS
XX
XX The present invention describes a method for generating a repository of
XX nucleic acids of tuf, fus, atp and/or recA genes from which probes
XX and/or primers are derived. The method comprises amplifying the nucleic
XX acids of determined algal, archaeal, bacterial, fungal and parasitical
XX species with a combination of defined primer pairs. The method can be
XX used for producing probes and/or primers for detecting one or more
XX related microorganisms e.g. algae, archaea, bacteria, fungi and
XX parasites, for universal detection and for specific and ubiquitous
XX detection and identification of an algal, archaeal, bacterial, fungal
XX and parasitical species, genus, family and group. A nucleic acid (I)
XX obtained using the method of the invention can be used for the universal
XX detection of any bacterium, fungus or parasite in a sample and for the
XX detection of at least one antimicrobial agent resistance gene or at
XX least one toxin gene. hexa nucleic acids are used for the specific and
XX ubiquitous detection and for identification of Streptococcus pneumoniae.
XX (I) can be used to design a therapeutic agent which is effective against
XX microorganisms. Microbial species or genus or family or phylum or group
XX which can be detected include Abiotrophia adiacens, Bordetella sp.,
XX Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
XX Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,

CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.

XX Sequence 1199 BP; 386 A; 286 C; 240 G; 287 T; 0 other;

Query Match 29.5%; Score 590; DB 22; Length 1199;
Best Local Similarity 99.7%; Pred. No. 4.1e-284;
Matches 690; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 761 CAAAACATCTGTGGATATTTTACAAATACAGACCAATACGTTGCCATATCCAGACGATGAA 820
Db 1 caaaacatctgtggtatttacaacagacgaatagctgtgctactccagcagatgaa 60
QY 821 TTGCAAGTCGCTTACCATTTGTTGATGTTTCTAACGGTAAAGTCATTTGCCAGCTAGGA 880
Db 61 ttgcaagtcgcttaccatgttgatgttcttaacggtlaaagtcattgcccagctagga 120
QY 881 GCACGGCATCAGTCAAGTAATGTTCTTGGATTAACAGACGATGAAGAAACACCGC 940
Db 121 gcaagcatcagtcagaagtaagttcccttcggaactaaccaagcagtagaacaacaccgc 180
QY 941 GACTGGGATCAACTATGAAACCGATCACAGACTATGCTCTGCTTGGAGTACGCTGTC 1000
Db 181 gactggggtacaaactatgaacccgacacagactatgctctgcttggtagtgc 240
QY 1001 TAGCATTCACATGCTACTATTCGTTTCACGATGAGCCCTATTAATCCTGGACAAATACT 1060
Db 241 tagcatccaactgctactatctgtcagatgagccctataactccctgggacaaataacc 300
QY 1061 CCTGTTTAACTAGTGGATGAGGCGCTACTTGGCAACATACCTTGCAATACGCCCTGCAA 1120
Db 301 ccgtttataacggtgagtgaggtcctacttggcaacatcaccttgcataagccctgcga 360
QY 1121 CAATCGGAAACGTCACGCGCTGGAAACCTTAACAAGGTGGAGCTCAACCGCGCCAG 1180
Db 361 caatcggaacgctcccgagcgtggaaccttaacaagtcgagaccacgcgcgaag 420
QY 1181 ACTTCTTAAATGCTCTAGCAATGACTACCCAGTATTTACTACTACTCAATGCCATTTC 1240
Db 421 acttcttaaatgctcctcgaatgactaccacgaatttcaactctcaaaatgacatttca 480
QY 1241 AGTACACAAACCGAATAGACAAATAATATGAGCAAGTAGTGAAGATGGCTGCT 1300
Db 481 agtaacacaaacgaatcagaacaaatataatgagcaagtagtgaagaagtcgctgcgtc 540
QY 1301 TACGCTGCTTTGCAAAATGTTGGAATCTACTATAACCAATGTATATCAATAAAGTCGC 1360
Db 541 tagcgtgcttggcaaatggtggaacttactataaccatgtatataccataaagtcgtc 600
QY 1361 TTTAGTATGGAGTGAAGAAAGAGTCTCTAATGTGCGAACTGCTGCCATGAAGAAACG 1420
Db 601 tttagtattggaatggaagaaagttctctaattgctggaactcgtgcattgaagaagacg 660
QY 1421 ACAGCCTATATGATGACGACGACATGATGAAG 1452
Db 661 acagcctatatgatgacgcagacatgatgaaac 692

RESULT 18

AAZ35945
AAZ35945 standard; DNA: 1260 BP.

AC AAZ35945;

DT 07-FEB-2000 (first entry)

DE Streptococcus pneumoniae pbp1a TER isolate g) nucleotide sequence.

XX Streptococcus pneumoniae; penicillin binding protein; pbp2b; pbp1a;
KW

KW transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
KM detection; identification; pneumococcal meningitis; ss.

XX Streptococcus pneumoniae.

XX 2A9807024-A.

XX 28-APR-1999.

XX 05-AUG-1998; 982A-0007024.

XX 01-AUG-1997; 972A-0006886.

XX (SAME-) SOUTH AFRICAN INST MEDICAL RES.

XX (UYWI-) UNIV WITWATERSRAND.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Klugman KP, Smith AM, Du Plessis M;

XX WPI; 1999-601770/51.

XX Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis

XX Claim 11, Fig 4; 63pp; English.

XX A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1a gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.

CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.

CC The assays can be adapted to detect other pathogens causing meningitis.

CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1a transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.

XX Sequence 1260 BP; 404 A; 281 C; 268 G; 307 T; 0 other;

Query Match 29.1%; Score 581; DB 20; Length 1260;
Best Local Similarity 99.7%; Pred. No. 1.8e-289;
Matches 681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 740 ACAATGTAGACCAAGAAGCTCAAAACATCTGGGATATTTTCAATACAGACGAATAC 799
Db 166 acaaatgtagaccagaagaagctcaaaacatctggtgatatattacaacagaagaatac 225
QY 800 GTTGCTATCCAGACGATGAATGCAATGCGCTTCTACCATTTGATGTTTCAACGCT 859
Db 226 gtgctcatccagacgataatgcaatgcaagtcgcttcaaccattgttgaatttcaacgct 285
QY 860 AAAGTCAATGCGCCAGTAGAGAGCGACGCATCAGCAAGTATGTTTCTCGAATTAAC 919
Db 286 aaagtcaatgcccagctagagcagccatcagccaagtaattgttcccttcggaattaac 345
QY 920 CAACGACTAGAAACACCGAGCTGGGATGCACTATGAACGATCACAGATATGCT 979
Db 346 caacgactagaaacacccgagctgggatactatgaacacgacacagactatgct 405
QY 980 CTTGCTTGGAGTACGGTGTCTACGATTCACATGCTACTATTCGTTCAAGATGACCCCTAT 1039
Db 406 ctgcttggagtagtcggtgtctacgatacctactactactcgttcaagatgagccctat 465
QY 1040 AACTACCTGGGACAAATACTCCTGTTTATTAATGAGGATAGGGGCTCTTTGGCAATC 1099
Db 466 aactaccctgggacaaataactcctgttataactggataggggctactcttggacaacatc 525

QY 1100 ACCTTGCAATACGCGCTGCAACATGCGAAGACGTCGAGCCGTGGAACCTTAACAG 1159
 |||||||
 Db 536 accttgcaataagccctgcgaacaatcggaacgctgcgtaaaccttaacaag 585
 QY 1160 GTGGGACTACCGCGCAAGACTTCTTAATGTCGTAGGAATGCAATCCAGTATT 1219
 |||||||
 Db 586 gtgggactacacgcgcgaagacttcttaattgctcgtgaatcgactccaagtatt 645
 QY 1220 CACTAGCAATGCGATTTCATGACACGCAATGACAAAATATGAGCAAGT 1279
 |||||||
 Db 646 cactactcaaatgacatttcaagtaacacacccgaatcgacaaaataatgagcaagt 705
 QY 1280 AGTGAAGAAGTGGCTGCTGCTTACGCTGCTTGAATGCTGGAATTACTTAACCA 1339
 |||||||
 Db 706 agtgaagaagatggctgctgcttaacgctgcttgcgaatggtggaacttaactaaca 765
 QY 1340 ATGTATATCATTAAGTCTGCTTACTGATGAGGATGAAAAGAGTTCTTAATGTCGGA 1399
 |||||||
 Db 766 atgtatatcatataaagctcgtcttaagtatgagtgagaaagagttcttaattcgcga 825
 QY 1400 ACTGTCGCATGAAGGAAGAC 1422
 |||||||
 Db 826 accgtgcacatgaggaagac 848

RESULT 19

AAH01026
 ID AAH01026 standard; DNA; 1229 BP.

AC AAH01026;

DT 24-JUL-2001 (first entry)

DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1017.

KW Species specific; genus specific; family specific; probe; detection;
 KW identification; algal; archaeal; bacterial; fungal; parasitical;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 KW vaccine; primer; ds.

OS Streptococcus pneumoniae.

PN MO200123604-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000MO-CA01150.

PR 28-SEP-1999; 99CA-2283458.

PR 19-MAY-2000; 2000CA-2307010.

PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.

PI Bergeron MC, Boissinot M, Huletsky A, Menard C, Ouellette M;

PI Picard FJ, Roy PH;

XX WPI: 2001-245006/25.

DR Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 PT archaeal, bacterial, fungal and parasitical species in a test sample -

PS Claim 27; Page 970-971; 1580pp; English.

CC The present invention describes a method for generating a repertoire of
 CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitical
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more

CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitical species, genus, family and group. A nucleic acid (I)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexA nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (I) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonas group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX

SO Sequence 1229 BP; 399 A; 287 C; 240 G; 303 T; 0 other;

Query Match 28.4%; Score 567; DB 22; Length 1229;
 Best Local Similarity 99.7%; Pred. No. 3.1e-282;
 Matches 667; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 745 TGTAGACCAAGAGCTCAAAACATCTGCGATTTTCAATACAGCAATACGTTGC 804
 |||||||
 Db 1 tgltagacaagaagcctcaaaaacatctgtggatattcaatacagcaatacgttcg 60
 QY 805 CTATCCAGAGATGAATTTGCAATGCGCTTCTACATTTGTGATGTTCTTAACGGTAAAGT 864
 |||||||
 Db 61 ctatccagaagatgaatttgcaatgcttctacattgttgatgcttcttaacgtaagt 120
 QY 865 CATTGCCAGCTAGAGACGCGCATGATCAATATGTTCTCTTGGAATTAACCAAGC 924
 |||||||
 Db 121 catlgccagctagtagagacgcacatgaagtaatgttctccttggaattacaaga 180
 QY 925 AGTAGAACAACACGGGAGCTGGGATCACTATGAACGCAATACAGATACGCTCTGTC 984
 |||||||
 Db 181 agtagaacaacaacgcgacgtgggagatcaactagaaacgcatcacagactatgccttc 240
 QY 985 CTTGAGTAGCGTGTCTAGCATTCACCTGCTACTATGTTCTACAGTAGAGCCCTTAATCTA 1044
 |||||||
 Db 241 ctggagtagcgtgtcttaagatcttaactatcgttcaatgaatgagagccataacta 300
 QY 1045 CCCTGGACAAATATCTCTGTTTATTAATCTGGATAGGGGCTACTTTGGCAATCACTCTT 1104
 |||||||
 Db 301 cccctggacaataatccctcgttataactatgagtagggctactcttgcaacatcacct 360
 QY 1105 GCAATACGCCCTGCAACATCCGGAACGTCGCCGCGGGAACCTTAACAAAGTCGG 1164
 |||||||
 Db 361 gcaatagccctgcaacaatcgcgaacgltccagccgttggaactcttaaaaggtcgg 420
 QY 1165 ACTCAACCGCGCCCAAGACTTCTTAATGCTGTAGCAATCGATACCAAGTATTCTA 1224
 |||||||
 Db 421 actcaacgcgcgccaagacttcttaaatggttcggaatcgatlaaccagaagttaacta 480
 QY 1225 CTGCAATGCCATTTCAGATGAACAACGAAATGACAAATAATGAGCAAGTAGTGA 1284
 |||||||
 Db 481 ctcaaatgcatctcaatgaatgaacacacgcatcagcaaaaataatgagcaagtagtga 540
 QY 1285 AAGATAGCGTGTGCTTACGCTGCTTGTGCAATGCTGGAACCTACTATTAACCAATGTA 1344
 |||||||
 Db 541 aagatagctgctgcttaacgctgcttgcgaatggttggaactactataaacaatgta 600
 QY 1345 TATCCATAAAGTGTCTTATGATGAGAGTGAAGAAAGATTCTTAATGTCGAAGTGC 1404
 |||||||
 Db 601 tatccataaagtgcttcttaagtagatggtggaaggttctcttaagtgcgaactcg 660
 QY 1405 TGCCATGAA 1413
 |||||||

Db 661 tgcgatgaa 669

RESULT 20
AA235943
ID AA235943 standard; DNA: 1260 BP.

XX
AC AA235943;
XX
DT 07-FEB-2000 (first entry)
XX

DE Streptococcus pneumoniae pbp1A TER isolate e) nucleotide sequence.
XX
KW Streptococcus pneumoniae; penicillin binding protein; pbp2b; pbp1A;
KW transpeptidase encoding region; TER: antibiotic resistance; diagnosis;
KW detection; identification; pneumococcal meningitis; ss.
XX
OS Streptococcus pneumoniae.
XX
PN 2A9807024-A.
XX
PD 28-APR-1999.
XX
PF 05-AUG-1998; 982A-0007024.
XX
PR 01-AUG-1997; 972A-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UYWI-) UNIT WITWATERSRAND.
PA (MED1-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
XX
DR WPI; 1999-601770/51.
XX

PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
XX useful for the diagnosis of pneumococcal meningitis
XX
PS Claim 11; Fig 4; 63pp; English.
XX

CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2b (pbp2b) gene
CC and the pbp1A gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1A transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX
SQ Sequence 1260 BP; 408 A; 286 C; 263 G; 303 T; 0 other;

Query Match 27.7%; Score 554; DB 20; Length 1260;
Best Local Similarity 99.4%; Pred. No.1.6e-275;
Matches 804; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 575 GAACAGTATGGAAGCAATACCAATTCATGATGGAGTACCAAGCTCAATFCA 634
DB 1 gaacagatcaggaagcaagcaatcctacacgaatcagcaagcctcaatca 60
QY 635 GCAAGTATTAACCTGCTTACATGATTAATTAATCAAGAGTCAATCAAGTTGAA 694
DB 61 gcaagtaattaccctcgtctacatgataattactcaagagcattcaatcaagttgaa 120
QY 695 GAAGAAACAGGCTATTAACCTACTACCAAGCTGGATGTCTACACAAATGTAGACCA 754

DB 121 gaagaaacagcgtataacacctcacaacactggaatgtaagctcaacaatgtagacaa 180
QY 755 GAAGCTCAAAAACATCTGTGGATATTACAAATACAGACGAATACGTTGCTATCCAGAC 814
DB 181 gaagctcaaaaacatcgtggaatattcaacatacagaacgaatcagctgctcaatccag 240
QY 815 GATGAATTTGCAAGTCGCTTACACATTTGATGATTTCTAACAGGTAAAGTCAATGCCAG 874
DB 241 gatgaattgcaagtcgctgctcaccatgttgatgtttcttaacggttaagtcattgccag 300
QY 875 CTAGAGCAGCCATCATGATCAATGTAATGTTCTTGGCAATTAACAGCAGTGAANA 934
DB 301 ctagagcagccatcagtcagtaagtaatttccttcggaatlaacccaagcagtagaana 360
QY 935 AACCGCAGCTGGGATCATCATATGAANCCGATCACAGCATGTGCTCGCTTGAGATAC 994
DB 361 aaccgcagctgggatacactaagaaacgaatcaacagctatgctcctctgagtlac 420
QY 995 GGTGCTACGATTCACATGCTACTATGCTTACAGATGAGCCCTTAATACCTGAGACA 1054
DB 421 ggtgctacgattcaactgctactatcgltcaagatgaaccttaactaaccctgggaca 480
QY 1055 AATACCTCCTTTTAACTGGGATAGGGGCTACTTTGGCAACATCACCTTGCAATACGCC 1114
DB 481 aataccctgttataactggaatagagggtactcttgacaacatcaccttgcaataagcc 540
QY 1115 CTGCAACAAATGGCAACGTCGCCAGCCGTGGAACACTTAACAGGTGCACTAACCCG 1174
DB 541 ctgcaacaatcgcgaacacgtccagccgtggaaccttaacaaagtgctgaacacgc 600
QY 1175 GCCAAGACTTTCCTAAATGCTAGAAATGCTAGAAATGCTAGAACCAATATCTACTCAATGCC 1234
DB 601 gccaaagacttccctaaatgctcgtgaaatcgacaccgaatctactactcaaatgctc 660
QY 1235 ATTTCAGTAACACACACGGAATCAGACAAAAAATATGAGCAAGTAGTGAAGAAATGCT 1294
DB 661 atttcagtaaacacacggaatcagacaaaaaataatggaagtagtgaagaatgctc 720
QY 1295 GCTGCTTACGCTGCTTTCGAAATGCTGGAATGCTAGTAATACCAATGTATATCATATA 1354
DB 721 gctgcttaacgctgcttcttgcgaatcgtggaacttaataaaccatgatatccataa 780
QY 1355 GTCGCTTTTACTGATGGAGTGAANAAGA 1383
DB 781 gtcgcttttactgattggtggaagtaanaaga 809

RESULT 21
AA296311
ID AA296311 standard; DNA: 2172 BP.
XX
AC AA296311;
XX
DT 10-APR-2000 (first entry)
XX
DE S. pneumoniae derived DNA from ORF #139.
XX
KW Treatment; prevention; disease; diagnosis; gene therapy; screening;
KW bacterial; antimicrobial; antibiotic; pathogenesis; infection; ss.
XX Streptococcus pneumoniae.
XX
PN WO9806734-A1.
XX
PD 19-FEB-1998.
XX
PF 15-AUG-1997; 97WO-US14436.
XX
PR 16-AUG-1996; 96US-0024022.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX

CC (1) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC can be detected include *Adiophora adiacens*, *Bordetella* sp.,
 CC *Corynebacterium* sp., *Enterobacteriaceae* group, *Escherichia coli*,
 CC *Mycobacteriaceae* family, *Pseudomonas* group, *Streptococcus* sp.,
 CC *Neisseria gonorrhoeae* and *Staphylococcus* sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX
 XX Sequence 2160 BP; 701 A; 494 C; 395 G; 570 T; 0 other;

Query Match 19.8%; Score 396; DB 22; Length 2160;
 Best Local Similarity 99.1%; Pred. No. 6.3e-194;
 Matches 696; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATATAAATCACTGATGCTGACTTGGGTTCTGAAGCGCGCTCAA 60
 DB 159 taaatctacgacataataaatcaatcattgctgacttggtctgaagcgcgctcaa 218

QY 61 TGCCCAAGCTAATGATATTCACAGATTTGGTTAAGCAATCGTTCTATCGAAGACA 120
 DB 219 tgcccaagctaatgatatccacagatttggtlaagcaatcgcttctatcgaagaca 278

QY 121 TCGCTTCTGACACAGAGGGGATTTGATACCATCGTATCCTGGAGACTTCTTGCCCAA 180
 DB 279 tcgcttctgacacaggggattgatacatccgtatcccttggaagcttcttgcgcaa 338

QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTCAACCACAGTTGATTAGTTGAC 240
 DB 339 tctgcaagcaatccctccaaggtgatacactcaccacaagattgataatctgac 398

QY 241 TTTACTTTCACTGACTTCGCGACGACGACTATTTCTCGTAAGGCTCAGGAAGTTGGTT 300
 DB 399 ttactttcaactcgacttcgcagcagacattctcgtlaagcttcagaagcttggtl 458

QY 301 AGCGATTCACTTGAACAAAAGCAACCAAGCAAGAATCTTGACTATATATAATAA 360
 DB 459 agcgattcagttgaacaaaagcaaccaaagaaatcttgactatataataa 518

QY 361 GGTTTACATGTTTGAAGGACTATGCAATGACAGACGAGCTAAAATCTATGTGTA 420
 DB 519 ggtttacatgttcaatgggcaactatgaaatgcagacagctcaaaactactatgtlaa 578

QY 421 AGACCTAATAATTTAGTTTACCTCAGTTAGCCTTGCTGGCTGGAATGCTCAGGAC 480
 DB 579 agacctaaataatttaagttactcctgtaagcttgcgtgctggaatgctccagacac 638

QY 481 AAACCAATATGACCCCTATTTCATCCAGAGACGCCCAAGACGCCGAATCTGTCTT 540
 DB 639 aaaccaatattgacccctatttcacatccagagaagcccaagacgcggaacttgctt 698

QY 541 ATCTGAATGAAAAATTAAGGCTATCTGCTGAACAGTATGAGAAGCAATCAATAC 600
 DB 699 atctgaatgaaaaatcaagttactcctgtaagcttgcgtgcaatgagaagcaagcaat 758

QY 601 ACCAATACGATGAGTACAAAGTTCAAATCGAATCGAAGTAATACCTGTTTCATGGA 660
 DB 759 accaataacgattgagtaaaaagttccaatcagcaagtaataacctgttactatgga 818

QY 661 TAAATTACCTCAGAGATCATCAATCAAGTTGAAGAAGAAC 702
 DB 819 taattacctcagaagatcatcaatcaagttgaaagaagaac 860

DT 24-JUL-2001 (first entry)
 XX Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2140.
 DE
 XX Species specific; genus specific; family specific; probe; detection;
 KW Identification; algal; archael; bacterial; fungal; parasitica;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 KW vaccine; primer; ds.
 XX
 XX Streptococcus pneumoniae.
 OS
 XX WO200123604-A2.
 PN
 XX 05-Apr-2001.
 PD
 XX 28-SEP-2000; 2000WO-CA01150.
 PF
 XX 28-SEP-1999; 99CA-2283458.
 PR 19-MAY-2000; 2000CA-2307010.
 PX
 XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
 PA
 XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;
 PI WPL; 2001-24506/25.
 DR
 XX Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 PT archael, bacterial, fungal and parasitica species in a test sample -
 PS Disclosure; Page 1508-1509; 1580pp; English.
 XX

XX The present invention describes a method for generating a repertory of
 CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archael, bacterial, fungal and parasitica
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archael, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archael, bacterial, fungal
 CC and parasitica species, genus, family and group. A nucleic acid (1)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexa nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of *Streptococcus pneumoniae*.
 CC (1) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include *Adiophora adiacens*, *Bordetella* sp.,
 CC *Corynebacterium* sp., *Enterobacteriaceae* group, *Escherichia coli*,
 CC *Mycobacteriaceae* family, *Pseudomonas* group, *Streptococcus* sp.,
 CC *Neisseria gonorrhoeae* and *Staphylococcus* sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX
 XX Sequence 2157 BP; 696 A; 494 C; 405 G; 562 T; 0 other;

Query Match 15.94%; Score 318; DB 22; Length 2157;
 Best Local Similarity 99.04%; Pred. No. 1.2e-153;
 Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

RESULT 23
 AAH02147
 ID AAH02147 standard; DNA: 2157 BP.
 AC
 AC AAH02147;
 XX
 XX

QY 1 TAAATCTACGACATATAAATCACTGATGCTGACTTGGGTTCTGAAGCGCGCTCAA 60
 DB 159 taaatctacgacataataaatcaatcattgctgacttggtctgaagcgcgctcaa 218

QY 61 TGCCCAAGCTAATGATATTCACAGATTTGGTTAAGCAATCGTTCTATCGAAGACA 120


```

DB 219 |tgcccaagctaatagatattcccaagatttggttaagcaatcgtttctatcgaagacca 278
QY 121 |tgcgtcttgcgacnccnaggggattgattacatccgtatccctggagcttcttggcnaa 180
DB 279 |tcgcttcttcgacaaaggagattgattccatccgtatccctggagcttcttggcnaa 338
QY 181 |tctgcaaaagcaattccctcgaagtgatcatcctcaccgaagttgattgaattgac 240
DB 339 |tctgcaaaagtaattccctcgaagtgatcatcctcaccgaagttgattgaattgac 398
QY 241 |ttacttttcaacttgcgttccgaccagactatttctgtaagctcaggaagcttggtt 300
DB 399 |ttacttttcaacttgcgttccgaccagactatttctgtaagctcaggaagcttggtt 458
QY 301 |acggattcgttagaacaacaaagcaacgaagaaattctgactactatataataa 360
DB 459 |agcgattcagtttagaacaacaaagcaacgaagaaattctgactactatataataa 518
QY 361 |ggtctacatgcttaattgggaactatggaatgcagacagcagctcaaaactatgtaa 420
DB 519 |ggtctacatgcttaattgggaactatggaatgcagacagcagctcaaaactatgtaa 578
QY 421 |agactcaataatttaagtttaacctgagttgacctgttgctggaatgctcagcacc 480
DB 579 |agactcaataatttaagtttaacctgagttgacctgttgctggaatgctcagcacc 638
QY 481 |aaaccaatttgacccttattcattccagacagcaccgacccgaacttggctt 540
DB 639 |aaaccaatttgacccttattcattccagacagcaccgacccgaacttggctt 698
QY 541 |atctgaataatgaacaaatcagcgtctgctgtaacagttgaaagcagctcaatc 600
DB 699 |atctgaataatgaacaaatcagcgtctgctgtaacagttgaaagcagctcaatc 758
QY 601 |accgaatttactgtagaactcaaaagttcaaatcagacagcttaattacccctgtaac 660
DB 759 |accgaatttactgtagaactcaaaagttcaaatcagacagcttaattacccctgtaac 818
QY 661 |taattacgtcgaaga 675
DB 819 |taattacgtcgaaga 833

```

RESULT 24

AAH01178 standard; DNA; 2160 BP.

AAH01178;

24-JUL-2001 (first entry)

Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1169.

Species specific; genus specific; family specific; probe; detection;
 identification; algal; archaeal; bacterial; fungal; parasitical;
 microorganism; diagnosis; translation elongation factor Tu; toxin;
 translation elongation factor G; RecA recombinase; resistance;
 catalytic subunit of proton-translocating ATPase; antimicrobial;
 vaccine; primer; ds.

Streptococcus pneumoniae.

WO200123604-A2.

05-APR-2001.

28-SEP-2000; 2000WO-CA01150.

28-SEP-1999; 99CA-2283458.

19-MAY-2000; 2000CA-2307010.

(INFE-) INFECTIO DIAGNOSTIC (IDI) INC.

XX Bergeron MC, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;
 XX WPI: 2001-245006/25.

PT Nucleic acid sequences are used to generate universal probes and
 PR primers which can be used to identify and detect the presence of algal,
 PT archaeal, bacterial, fungal and parasitical species in a test sample -

PS Disclosure: Page 1044-1045; 1580pp; English.

The present invention describes a method for generating a repository of
 CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitical
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitical species, genus, family and group. A nucleic acid (I)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexA nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (I) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.

SO Sequence 2160 BP; 698 A; 494 C; 403 G; 565 T; 0 other;

Query Match 15.9%; Score 318; DB 22; Length 2160;
 Best Local Similarity 99.0%; Pred. No. 1.2e-153;
 Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 1 TAAATCTACGACATTAATAATCACTGATTGCTGACTTGGGTTTGAACGCCGCTCA 60
DB 159 taaatctacgaacaataaaatcaactatgctgacttgggttctgaaacgcgcgtcaa 218
QY 61 TGCCCAAGCTAATGATATTCCCAAGATTGGTTAAGGCAATCGTTTATGGAACGCA 120
DB 219 tgcccaagctaatgatattcccaagattggtttaaggaacatcgtttccatcgaagacca 278
QY 121 TCGCTTCTTCGACACAGGGGGATTGATCCATCCGTATCTCGGAGCTTTCTTGGCA 180
DB 279 tcgcttcttcgacacaaaggagattgattccatccgtatccctggagcttcttggcnaa 338
QY 181 TCTGCAAGCAATTCCCTCAGAGTGATCAACTCTCACCCAAAGTTTAAATTGATGAC 240
DB 339 tctgcaaaagtaattccctcgaagtgatcatcctcaccgaagttgattgaattgac 398
QY 241 TTACTTTTCAACTTGCAGCTTCGACAGACTATTCTGTAAGGGCTCAGGAAGCTTGTT 300
DB 399 ttacttttcaacttgcgttccgaccagactatttctgtaagctcaggaagcttggtt 458
QY 301 ACCGATTGCTTAGAACAACAAAGCAACGACGAAGAAATCTTGACTACTATATAATTA 360
DB 459 agcgattcagtttagaacaacaaagcaacgaagaaattctgactactatataataa 518
QY 361 GGTCTACATGTTAATGGAACTATGGAATGCAAGACAGCTCAAAACTACTATGTTAA 420
DB 519 ggtctacatgcttaattgggaactatggaatgcagacagcagctcaaaactatgtaa 578

```


QY 421 AGACCTCAATATTTAGTTTACCTCAGTTAGCCTTGCTGGTGAATGCTCAGCACC 480
CC |
CC |
CC |
Db 579 agaccctcaataatlaattagcttagcttagcttgctgctggaatgctcagcagacc 638
QY 481 AAACCAATATGACCCCTATTCATCCAGAGAGAGCCCAAGACCCGGCAACTTGCTT 540
CC |
CC |
CC |
Db 639 aaaccataatgaccccatcacatccagaagcagcccaagccgcgaaactggtcct 698
QY 541 ATGTGAATGAAAAATCAAGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC 600
CC |
CC |
CC |
Db 699 atctgaaatgaaaaaatcaaggttacatctctgtcgaaagcagtaagaaagcagtaac 758
QY 601 ACCAATTACTGATGAGTACAAAGTCTCAATGACGCAAGTAATTCCTCTTACATGGA 660
CC |
CC |
CC |
Db 759 accaattactgactgactacaagtcacaaatcagcaagtaattacctgcttacatgga 818
QY 661 TAATTACTCTAAGGA 675
CC |
CC |
CC |
Db 819 taattactcaagga 833
RESULT 25
AAH02056
ID AAH02056 standard; DNA; 2160 BP.
XX
XX AAH02056;
XX
XX 24-JUL-2001 (first entry)
XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2049.
XX
XX Species specific; genus specific; family specific; probe; detection;
XX identification; algal; archaeal; bacterial; fungal; parasitical;
XX microorganism; diagnosis; translation elongation factor Tu; toxin;
XX translation elongation factor G; RecA recombinase; resistance;
XX catalytic subunit of proton-translocating ATPase; antimicrobial;
XX vaccine; primer; ds.
XX
XX Streptococcus pneumoniae.
XX
XX WO200123604-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-CA01150.
XX
XX 28-SEP-1999; 99CA-2283458.
XX
XX 19-MAY-2000; 2000CA-2307010.
XX
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
XX Picard FJ, Roy PH;
XX
XX WPI; 2001-245006/25.
XX
XX
XX Nucleic acid sequences are used to generate universal probes and
XX primers which can be used to identify and detect the presence of algal,
XX archaeal, bacterial, fungal and parasitical species in a test sample -
XX
XX
XX Disclosure; Page 1464-1465; 1580pp; English.
XX
XX
XX
XX The present invention describes a method for generating a repertory of
XX nucleic acids of tuf, fus, atpd and/or recA genes from which probes
XX and/or primers are derived. The method comprises amplifying the nucleic
XX acids of determined algal, archaeal, bacterial, fungal and parasitical
XX species with a combination of defined primer pairs. The method can be
XX used for producing probes and/or primers for detecting one or more
XX related microorganisms e.g. algae, archaea, bacteria, fungi and
XX parasites, for universal detection and for specific and ubiquitous
XX detection and identification of an algal, archaeal, bacterial, fungal
XX and parasitical species, genus, family and group. A nucleic acid (I)
XX obtained using the method of the invention can be used for the universal

CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexA nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (I) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
XX
XX Sequence 2160 BP; 696 A; 495 C; 402 G; 567 T; 0 other;

Query Match 15.9%; Score 318; DB 22; Length 2160;
Best Local Similarity 99.0%; Pred. No. 1.2e-153;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAAATCTAGACAAATAAATCAACTCATCTGCTGAGTTGGTTCGAAGCCGCGTCAA 60
Db 159 taaatctagacaataaataaatacaactcatctgctgagttggttcgaagccgcgtcaa 218
QY 61 TGCCCAAGCTAATGATATTCCTCCACAGATTGGTTAAGGCAATGCTTCTATGGAAGCA 120
Db 219 tgcccaagctaatgatatctccacagattggtttaaggcaatgcttctatggaagca 278
QY 121 TCGCTTTTGACACACAGGGGATTTGATACCTCCGATCTCTGGAGCTTTCTTGCCAA 180
Db 279 tcgcttttgacacacaggggatTTGATACCTCCGATCTCTGGAGCTTTCTTGCCAA 338
QY 181 TCTGCAAGCAATTCCTCCAGGTGATCACTCTCAACCAAGTGTGATTAAGTTGAC 240
Db 339 tctgcaagcaatTCCTCCAGGTGATCACTCTCAACCAAGTGTGATTAAGTTGAC 398
QY 241 TTACTTTCAACTTCGACTTCGACGAGACTATTTCTGTAAGGCTCAGGAAGCTTGTT 300
Db 399 ttaactttcaactTCGACTTCGACGAGACTATTTCTGTAAGGCTCAGGAAGCTTGTT 458
QY 301 AGCGATTCACTTACACAAAAGACACCAAGCAAGAAATCTTGACTCTATATTAATTA 360
Db 459 agcgattcaactTACACAAAAGACACCAAGCAAGAAATCTTGACTCTATATTAATTA 518
QY 361 GGCTACATGCTAATGGAAGTATGGAATGACAGACAGCTCAAAACTATGATGTTAA 420
Db 519 ggctacatgctAATGGAAGTATGGAATGACAGACAGCTCAAAACTATGATGTTAA 578
QY 421 AGACCTCAATATTTAGTTTACCTCAGTTAGCCTTGCTGGTGAATGCTCAGCACC 480
Db 579 agaccctcaataatlaattagcttagcttagcttgctgctggaatgctcagcagacc 638
QY 481 AAACCAATATGACCCCTATTCATCCAGAGAGAGCCCAAGACCCGGCAACTTGCTT 540
Db 639 aaaccataatgaccccatcacatccagaagcagcccaagccgcgaaactggtcct 698
QY 541 ATGTGAATGAAAAATCAAGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC 600
Db 699 atctgaaatgaaaaaatcaaggttacatctctgtcgaaagcagtaagaaagcagtaac 758
QY 601 ACCAATTACTGATGAGTACAAAGTCTCAATGACGCAAGTAATTCCTCTTACATGGA 660
Db 759 accaattactgactgactacaagtcacaaatcagcaagtaattacctgcttacatgga 818
QY 661 TAATTACTCTAAGGA 675
Db 819 taattactcaagga 833

RESULT 26
AA235946

ID AA235946 standard; DNA: 1260 BP.
XX
AC AA235946;
XX
DT 07-FEB-2000 (first entry)
XX
DE Streptococcus pneumoniae pbp1A TER isolate h) nucleotide sequence.
XX
KM Streptococcus pneumoniae; penicillin binding protein: pbp2B; pbp1A;
KM transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
KM detection; identification; pneumococcal meningitis; ss.
XX
OS Streptococcus pneumoniae.
XX
PN ZA9807024-A.
XX
PD 28-APR-1999.
XX
PF 05-AUG-1998; 982A-0007024.
XX
PR 01-AUG-1997; 97ZA-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UYWI-) UNIT WITWATERSRAND.
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
XX
DR WPI; 1999-601770/51.
XX
PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
XX useful for the diagnosis of pneumococcal meningitis
XX
PS Claim 11; Fig 4; 63pp: English.
XX
CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1A gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1A transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX
SQ Sequence 1260 BP; 406 A; 284 C; 267 G; 303 T; 0 other;

Query Match 13.3%; Score 266; DB 20; Length 1260;
Best Local Similarity 99.1%; Pred. No. 8.3e-127;
Matches 466; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 575 GAACAGTATGAGCAAGCAGTACATACCAATTAATGATGAGCTACCAAGTCTCAATCA 634
DB 1 gaacagtatgagaagacgatacaatacaatactgactgatacaaaagctcaaatca 60
QY 635 GCAGGTAATTAACCTGCTTACATGATGAATTAATTAATCAAGCAATCAAGTTGAA 694
DB 61 gcaagtaattaccctgcttaccatgataatcaccagaagatcacaatcaagttgaa 120
QY 695 GAAGAACAAGGCTATTAACCTACTACACACTGGATGGATGCTATACCAAAATGTAGACCA 754
DB 121 gaagaacaagagataaacctgcacacactggatgatagtctcaacaatgtagaccaa 180
QY 755 GAAGCTCAAAAACATCTGTGGGATATTATACAAATACAGAGCAATAGCTGTCCAGAC 814
|||||

DB 181 gaagctcaaaaacatctgtggatattacaatacagacgaatcgttgcctacagac 240
QY 815 GATGAATTCGACAGTCGCTTACCATTTGATGATTTCTAAGCGTAAAGCATTTGCCAG 874
DB 241 gatgaattcgaagctgcttaccatgttgatgttcttaacggttaagatcattgccag 300
QY 875 CTAGAGACACGCCATCAGTCAAGTAATGTTCTTCGGAATTAACCAAGCATAGAAACA 934
DB 301 ctgagagcacgcatacgaatgaatgatttcttcgtgaattacaagaagaggaacaa 360
QY 935 AACCGCAGCTGGGATCACTATGAACGCATCACAGACTATGCTCTGCTTGAGCTAC 994
DB 361 aaccgcagctgggatacactatgaaacgcatacagactatctcctgcttgaatac 420
QY 995 GAGTCTACGATCACTGCTACTATAGCTTCAGATGAGCCCTATACTA 1044
DB 421 ggtctactagatcaactgactactgcttccagatgagccctatacta 470

RESULT 27
AAH02059
ID AAH02059 standard; DNA: 1195 BP.
XX
AC AAH02059;
XX
DT 24-JUL-2001 (first entry)
XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2052.
XX
XX Species specific; genus specific; family specific; probe; detection;
XX identification; algal; archaeal; bacterial; fungal; parasitical;
XX microorganism; diagnosis; translation elongation factor Tu; toxin;
XX translation elongation factor G; RecA recombinase; resistance;
XX catalytic subunit of proton-translocating ATPase; antimicrobial;
XX vaccine; primer; ds.
XX
OS Streptococcus pneumoniae.
XX
PN MO200123604-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-CA01150.
XX
PR 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX
PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
XX WPI; 2001-245006/25.
XX
DR

Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitological species in a test sample -
PS Disclosure; Page 1467; 1580pp; English.

The present invention describes a method for generating a repository of
CC nucleic acids of tuf, fus, atpd and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitological
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archaeal, bacterial, fungal
CC and parasitological species, genus, family and group. A nucleic acid (I)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at

PF 28-SEP-2000; 2000WO-CA01150.
XX
PR 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX
XX (INFE-) INFECTION DIAGNOSTIC (IDI) INC.
PI Bergeron MG, Bolssnot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX WPI; 2001-245006/25.
XX
XX Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitological species in a test sample -
XX
XX Claim 27; Page 1470; 1580pp; English.
XX
XX The present invention describes a method for generating a repository of
CC nucleic acids of tuf, fus, atpd and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitological
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archaeal, bacterial, fungal
CC and parasitological species, genus, family and group. A nucleic acid (I)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexa nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (I) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
XX Sequence 1212 BP; 381 A; 292 C; 244 G; 295 T; 0 other;
SO

Query Match 8.9%; Score 178; DB 22; Length 1212;
Best Local Similarity 99.6%; Pred. No. 2.3e-81;
Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1739 GATTGGAATATACCAAGGGGCTCTACAGAAATGGAATTCGATTTAAATGGTGGCT 1798
DB 984 gattggaatataccagaaggggctctacagaatggaatctgatttaaaatggtgct 1043
OY 1799 CATTACAGTGAACACCTGCTCCACAAACCCCTCAACTCAAGTTCACAGCTCA 1858
DB 1044 cgttctacgttgagctcactcgtctccacaacaccccccatcaactgaaatgctcaagctca 1103
OY 1859 TCATCAGATAGTTCACATTCACAGTTCAGCTCAACCACTCCAGACCAATAATAGTACG 1918
DB 1104 tcatcagatagttcaactcactcagcttagctcaaccactccaagcaataatagtagc 1163
OY 1919 ACTACCAATCTCAACATATACGCAACATCAATCAATCAATCAATCAATCAATCAATCA 1967
DB 1164 actaccatctcacaataataatagcaacaataataataataataataataataataata 1212

RESULT 30
AAH01025
ID AAH01025 standard; DNA: 1222 BP.
XX
AC AAH01025;

XX 24-JUL-2001 (first entry)
DT
XX Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1016.
DE
XX
XX Species specific; genus specific; family specific; probe; detection;
KW identification; algal; archaeal; bacterial; fungal; parasitological;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds;
XX
XX Streptococcus pneumoniae.
OS
XX WO200123604-A2.
PN
XX
XX 05-APR-2001.
PD
XX
XX 28-SEP-2000; 2000WO-CA01150.
PF
XX
XX 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX
XX (INFE-) INFECTION DIAGNOSTIC (IDI) INC.
PI Bergeron MG, Bolssnot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX WPI; 2001-245006/25.
XX
XX Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitological species in a test sample -
XX
XX Claim 27; Page 969-970; 1580pp; English.
XX
XX The present invention describes a method for generating a repository of
CC nucleic acids of tuf, fus, atpd and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitological
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archaeal, bacterial, fungal
CC and parasitological species, genus, family and group. A nucleic acid (I)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexa nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (I) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
XX Sequence 1222 BP; 389 A; 270 C; 233 G; 330 T; 0 other;
SO

Query Match 8.9%; Score 178; DB 22; Length 1222;
Best Local Similarity 99.6%; Pred. No. 2.3e-81;
Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1739 GATTGGAATATACCAAGGGGCTCTACAGAAATGGAATTCGATTTAAATGGTGGCT 1798
DB 994 gattggaatataccagaaggggctctacagaatggaatctgatttaaaatggtgct 1053

QY 1799 GATTTCAGTGGAACTGACCTGCTCTCACACACACCCCATCACTGAAAGTTCAAGCTCA 1858
|||||
Db 1054 cgttctacgttgagctcaccgcgtccacaacaccccatcaactgaaagttcaagctca 1113
QY 1859 TCATCAGATAGTTCACACTTTCACACTGACACACACTCCACACACAAATATAGTACG 1918
|||||
Db 1114 tcatcagatagttcaacttcacagctcagctcaaccactccaagcacaataatagtagcg 1173
QY 1919 ACTACCAATCCCTCAATATACGCAACAAATCAATACACCCCTGATC 1967
|||||
Db 1174 actaccatccttaacataatagcgaacatcaataacacccctgac 1222

RESULT 31
AAH01022
ID AAH01022 standard; DNA: 1220 BP.
XX
AC AAH01022;
XX
XX 24-JUL-2001 (first entry)
XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1013.
XX
KW Species specific; genus specific; family specific; probe; detection;
KW Identification; algal; archaeal; bacterial; fungal; parasitica;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
OS Streptococcus pneumoniae.
XX
PN WO260123604-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-CA01150.
XX
PR 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX
PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
DR WPI: 2001-245006/25.
XX
PT Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitica species in a test sample -
PS Claim 27; Page 967-968; 1580pp; English.
XX

The present invention describes a method for generating a repertory of nucleic acids of tuf, fus, atpd and/or recA genes from which probes and/or primers are derived. The method comprises amplifying the nucleic acids of determined algal, archaeal, bacterial, fungal and parasitica species with a combination of defined primer pairs. The method can be used for producing probes and/or primers for detecting one or more related microorganisms e.g. algae, archaea, bacteria, fungi and parasites, for universal detection and for specific and ubiquitous detection and identification of an algal, archaeal, bacterial, fungal and parasitica species, genus, family and group. A nucleic acid (II) obtained using the method of the invention can be used for the universal detection of any bacterium, fungus or parasite in a sample and for the detection of at least one antimicrobial agent resistance gene or at least one toxin gene. hexA nucleic acids are used for the specific and ubiquitous detection and for identification of Streptococcus pneumoniae. (I) can be used to design a therapeutic agent which is effective against microorganisms. Microbial species or genus or family or phylum or group which can be detected include Adenoviridae, Bordetella sp., Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,

CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp., using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
SQ Sequence 1220 BP: 388 A; 269 C; 233 G; 330 T; 0 other;

Query Match 8.8%; Score 175; DB 22; Length 1220;
Best Local Similarity 99.6%; Pred. No. 8.1e-80;
Matches 225; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1739 GATTTCAGTGGAACTGACCTGCTCTCACACACACCCCATCACTGAAAGTTCAAGCTCA 1858
|||||
Db 995 gatgggaataatataccagagggggtctcacaagaatggaatcgtatataaaggggct 1054
QY 1799 GATTTCAGTGGAACTGACCTGCTCTCACACACACCCCATCACTGAAAGTTCAAGCTCA 1858
|||||
Db 1055 cgttctacgttgagctcaccgcgtccacaacaccccatcaactgaaagttcaagctca 1114
QY 1859 TCATCAGATAGTTCACACTTTCACACTGACACACACTCCACACACAAATATAGTACG 1918
|||||
Db 1115 tcatcagatagttcaacttcacagctcagctcaaccactccaagcacaataatagtagcg 1174
QY 1919 ACTACCAATCCCTCAATATACGCAACAAATCAATACACCCCTG 1964
|||||
Db 1175 actaccatccttaacataatagcgaacatcaataacacccctg 1220

RESULT 32
AAH02069
ID AAH02069 standard; DNA: 1216 BP.
XX
AC AAH02069;
XX
XX 24-JUL-2001 (first entry)
XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2062.
XX
KW Species specific; genus specific; family specific; probe; detection;
KW Identification; algal; archaeal; bacterial; fungal; parasitica;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
OS Streptococcus pneumoniae.
XX
PN WO200123604-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-CA01150.
XX
PR 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX
PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
DR WPI: 2001-245006/25.
XX
PT Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitica species in a test sample -
PS Claim 27; Page 1474-1475; 1580pp; English.
XX

The present invention describes a method for generating a repertory of

KW	catalytic subunit of proton-translocating ATPase; antimicrobial;
KW	vaccine; primer; ds.
XX	
OS	Streptococcus pneumoniae.
PM	MO200123604-A2.
XX	
PD	05-APR-2001.
XX	
PF	28-SEP-2000; 2000MO-CA01150.
XX	
PR	28-SEP-1999; 99CA-2283458.
XX	
PR	19-MAY-2000; 2000CA-2307010.
XX	
PA	(INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX	
PI	Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI	Picard FJ, Roy PH;
XX	
XX	WPI; 2001-245006/25.
XX	
XX	Nucleic acid sequences are used to generate universal probes and
PT	primers which can be used to identify and detect the presence of algal,
PT	archaeal, bacterial, fungal and parasitical species in a test sample -
XX	
PS	Claim 27; Page 961-962; 1580pp; English.
XX	
CC	The present invention describes a method for generating a repertory of
CC	nucleic acids of tuf, fus, atpd and/or recA genes from which probes
CC	and/or primers are derived. The method comprises amplifying the nucleic
CC	acids of determined algal, archaeal, bacterial, fungal and parasitical
CC	species with a combination of defined primer pairs. The method can be
CC	used for producing probes and/or primers for detecting one or more
CC	related microorganisms e.g. algae, archaea, bacteria, fungi and
CC	parasites, for universal detection and for specific and ubiquitous
CC	detection and identification of an algal, archaeal, bacterial, fungal
CC	and parasitical species, genus, family and group. A nucleic acid (I)
CC	obtained using the method of the invention can be used for the universal
CC	detection of any bacterium, fungus or parasite in a sample and for the
CC	detection of at least one antimicrobial agent resistance gene or at
CC	least one toxin gene. hexa nucleic acids are used for the specific and
CC	ubiquitous detection and for identification of Streptococcus pneumoniae.
CC	(I) can be used to design a therapeutic agent which is effective against
CC	microorganisms. Microbial species or genus or family or phylum or group
CC	which can be detected include Alotrophia adiacens, Bordetella sp.,
CC	Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC	Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC	Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC	provides faster results than substrate specificity tests as results can
CC	be determined in an hour and improved accuracy is also achieved.
CC	AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC	which are given in the exemplification of the present invention.
XX	
XX	Sequence 1212 BP; 391 A; 265 C; 235 G; 321 T; 0 other;
XX	
Query Match	6.4%; Score 127; DB 22; Length 1212;
Best Local Similarity	99.1%; Pred. No. 4, 9e-55;
Matches 227; Conservative 0; Mismatches 2; Indels 0; Gaps	
OY	1739 GATTGGAAATATACGAGAGGGGCTCTACAGAAATGAGAAATTCATTATTAATAAATGGTGCT 1798
DB	
DB	984 gatggaataataccagaggggctctacagaatgagaattcgatttaaaatggtgct 1043
OY	1799 GGTTCCTACGTGGAACTACCTGCTCCACAAACACCCCATCAACTGAAAGTTTAACCTCA 1858
DB	
DB	1044 cgtcttaagtgagctcaacctcgcacacaacaccccatcaactgaaagttaagctca 1103
OY	1859 TCATCAGATTAATTCACCTTCACAGCTTACAGTCAACCACTCCAGGACCAATTAATAGTAGC 1918
DB	
DB	1104 tcatcagatagttcaacttcaacgcttcaactcaaacctctcaagcacaataatagtagc 1163
OY	1919 ACTACCAATTCCTTAACATTAATACGACCAATCAATTAACCAACCCCTGATC 1967

Db 1164 actaccatcttaacataatagcgaacatcaataacccctgac 1212

|||||

RESULT 35

AAH01017

ID AAH01017 standard; DNA: 1223 BP.

XX AAH01017;

XX 24-JUL-2001 (first entry)

DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1008.

XX

XX Species specific: genus specific; family specific; probe: detection:

KW identification: algal; archaeal; bacterial; fungal; parasitical;

KW microorganism; diagnosis; translation elongation factor Tu; toxin;

KW translation elongation factor G; RecA recombinase; resistance;

KW catalytic subunit of proton-translocating ATPase; antimicrobial;

KW vaccine; primer; ds.

XX

OS Streptococcus pneumoniae.

PN WO200123604-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000MO-CA01150.

PR 28-SEP-1999; 99CA-2283458.

PR 19-MAY-2000; 2000CA-2307010.

XX

PA (INFE-) INFECTION DIAGNOSTIC (IDI) INC.

XX

XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;

PI Picard FJ, Roy PH;

DR WPI: 2001-245006/25.

XX

XX Nucleic acid sequences are used to generate universal probes and

XX primers which can be used to identify and detect the presence of algal,

PT archael, bacterial, fungal and parasitcal species in a test sample -

XX

PS Claim 27: Page 964: 1580pp; English.

XX

CC The present invention describes a method for generating a repertory of

CC nucleic acids of tuf, fus, atpd and/or recA genes from which probes

CC and/or primers are derived. The method comprises amplifying the nucleic

CC acids of determined algal, archael, bacterial, fungal and parasitcal

CC species with a combination of defined primer pairs. The method can be

CC used for producing probes and/or primers for detecting one or more

CC related microorganisms e.g. algae, archaea, bacteria, fungi and

CC parasites. for universal detection and for specific and ubiquitous

CC detection and identification of an algal, archael, bacterial, fungal

CC and parasitcal species, genus, family and group. A nucleic acid (1)

CC obtained using the method of the invention can be used for the universal

CC detection of at least one antimicrobial agent resistance gene or at

CC least one toxin gene. hexa nucleic acids are used for the specific and

CC ubiquitous detection and for identification of Streptococcus pneumoniae.

CC (1) can be used to design a therapeutic agent which is effective against

CC microorganisms. Microbial species or genus or family or phylum or group

CC which can be detected include Alotrophia adiacens, Bordetella sp.,

CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,

CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,

CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests

CC provides faster results than substrate specificity tests as results can

CC be determined in an hour and improved accuracy is also achieved.

CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes

CC which are given in the exemplification of the present invention.

XX

SO Sequence 1223 BP: 394 A: 271 C: 235 G: 323 T: 0 other:

PT archaeal, bacterial, fungal and parasitical species in a test sample -
 XX
 PS Claim 27; Page 963; 1580pp; English.
 XX
 CC The present invention describes a method for generating a repository of
 CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitical
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitical species, genus, family and group. A nucleic acid (I)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexa nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (1) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX
 SQ Sequence 1218 BP; 392 A; 271 C; 233 G; 322 T; 0 other;

Query Match 6.3%; Score 126; DB 22; Length 1218;
 Best Local Similarity 100.0%; Pred. No. 1.6e-54;
 Matches 126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1842 CTGAAGTTCAGCTCATCATGATGTTCACTTACAGCTTACAGCTTCAACCACTCCAA 1901
 Db 1093 ctgaagttcagctcatcatgatttcaacttcacagctcagctcaaccactccaa 1152
 QY 1902 GCACAAATTAATAGTACGACTGCATCTTAATACCAATTAATACCAATTAATACCA 1961
 Db 1153 gcacaataaagtagcagctacaccccaataaataacgaacaatacaataacc 1212
 QY 1962 CTGATC 1967
 Db 1213 ctgctc 1218

RESULT 38
 AAH01015
 ID AAH01015 standard; DNA: 1213 BP.

AC AAH01015;
 DT 24-JUL-2001 (first entry)
 DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1006.
 XX
 XX Species specific; genus specific; family specific; probe; detection;
 KW identification; algal; archaeal; bacterial; fungal; parasitical;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 KW vaccine; primer; ds.
 XX
 XX Streptococcus pneumoniae.

XX
 PN WO200123604-A2.
 XX
 PD 05-APR-2001.
 XX

PF 28-SEP-2000; 2000WO-CA01150.
 XX
 PR 28-SEP-1999; 99CA-2283458.
 PR 19-MAY-2000; 2000CA-2307010.
 XX
 XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
 PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;
 XX
 DR WPI; 2001-24506/25.
 XX
 XX
 PT Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 PT archaeal, bacterial, fungal and parasitical species in a test sample -
 PS Claim 27; Page 962-963; 1580pp; English.

CC The present invention describes a method for generating a repository of
 CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitical
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitical species, genus, family and group. A nucleic acid (I)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexa nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (1) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX
 SQ Sequence 1213 BP; 393 A; 266 C; 235 G; 319 T; 0 other;

Query Match 6.1%; Score 122; DB 22; Length 1213;
 Best Local Similarity 99.1%; Pred. No. 1.9e-52;
 Matches 222; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1739 GATTGGAATATATACAGAGGGCTCTACAGAAATGAGAAATCGTATTTAAAAATGGTCT 1798
 Db 990 gattggaatatacagagggctctacagaaatgagaaatcgtaattaaatgggtct 1049
 QY 1799 CGTTCTAGTGTGAACTGCTGCTCCACACACACCCCATCATCACTGAAAGTTCAAGCTCA 1858
 Db 1050 cgttctagtgtagtgcctccgctccacaacaaccccatcaactgaagttcaagctca 1109
 QY 1859 TCATCAGATAGTCAACTTCACTGCTTACAGTCAACCACTCCACACACAAATATAGTAGTCA 1918
 Db 1110 tcatacagtagtcaacttcacagcttcaactcaaccactccaagacaataatagtagc 1169
 QY 1919 ACTACCAATCTCAACATTAATACGACACCAATCAATATACACCC 1962
 Db 1170 actacaatacctcaacaataaataacgaacaatacaatacaacc 1213

RESULT 39
 AAH01018
 ID AAH01018 standard; DNA: 1214 BP.
 XX
 AC AAH01018;

XX 24-JUL-2001 (first entry)
 XX
 DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1009.
 XX
 KW Species specific; genus specific; family specific; probe; detection;
 KW identification; algal; archaeal; bacterial; fungal; parasitica;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 KW vaccine; primer; ds.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200123604-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-CA01150.
 XX
 PR 28-SEP-1999; 99CA-2283458.
 XX 19-MAY-2000; 2000CA-2307010.
 XX
 PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
 XX
 PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;
 XX
 DR WPI: 2001-245006/25.
 XX
 PT Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 PT archaeal, bacterial, fungal and parasitica species in a test sample -
 XX
 Claim 27; Page 964-965; 1580pp; English.
 XX
 PS The present invention describes a method for generating a repertory of
 XX nucleic acids of tuf, fus, atp and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitica
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitica species, genus, family and group. A nucleic acid (I)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexA nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (I) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC that are given in the exemplification of the present invention.
 XX
 SQ Sequence 1214 BP; 391 A; 270 C; 232 G; 321 T; 0 other;

Query Match 6.1%; Score 122; DB 22; Length 1214;
 Best Local Similarity 100.0%; Pred. No. 1.9e-52;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1842 CTGAAGTTCACGCTATCATGATGATTCACCTTCACAGCTTACTCTCAACACCTCCAA 1901
 Db 1093 ctgaagttcaagctatcatcatagcttcaacttcaagctatgctcaacacccca 1152

OY 1902 GCACAATATAGTACGACTACCATCTTACACATATAGCACAATACCAACCC 1961
 Db 1153 gcaacaataatagtagactaccatccatccataacataatagcacaataacaccc 1212
 OY 1962 CT 1963
 Db 1213 ct 1214
 RESULT 40
 AAH01020
 ID AAH01020 standard; DNA; 1207 BP.
 XX
 AC AAH01020;
 XX
 DT 24-JUL-2001 (first entry)
 XX
 DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1011.
 XX
 KW Species specific; genus specific; family specific; probe; detection;
 KW identification; algal; archaeal; bacterial; fungal; parasitica;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 KW vaccine; primer; ds.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200123604-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-CA01150.
 XX
 PR 28-SEP-1999; 99CA-2283458.
 XX 19-MAY-2000; 2000CA-2307010.
 XX
 PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
 XX
 PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;
 XX
 DR WPI: 2001-245006/25.
 XX
 PT Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 PT archaeal, bacterial, fungal and parasitica species in a test sample -
 XX
 Claim 27; Page 966; 1580pp; English.
 XX
 PS The present invention describes a method for generating a repertory of
 XX nucleic acids of tuf, fus, atp and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitica
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitica species, genus, family and group. A nucleic acid (I)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexA nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (I) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.

CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
SQ Sequence 1207 BP; 388 A; 268 C; 231 G; 320 T; 0 other;

Query Match 6.1%; Score 121; DB 22; Length 1207;
Best Local Similarity 100.0%; Pred. No. 6.2e-52;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1842 CTGAAAGTTCAAGCTCATCATGTTCACTTCAAGTCTAGCTCAACCACTCCAA 1901
|
DB 1087 ctgaaagttcaagctcatcatcagatagttcaacttcacagctcagctcaaccactccaa 1146
QY 1902 GCACAATTAATAGTACGACGTACCAATCTTAACATAATAGCAGCAATATCAATCAACCC 1961
|
DB 1147 gcacaataataagtagcagactaccatcctaataataagcaacaatacaatacaacc 1206
QY 1962 C 1962
DB 1207 c 1207

RESULT 41

AAH01013
ID AAH01013 standard; DNA; 1212 BP.

AAH01013;

24-JUL-2001 (first entry)

Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1004.

Species specific; genus specific; family specific; probe: detection;
KW identification: algal; archaeal; bacterial; fungal; parasitica;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.

Streptococcus pneumoniae.

WO200123604-A2.

05-APR-2001.

28-SEP-2000; 2000WO-CA01150.

28-SEP-1999; 99CA-2283458.

19-MAY-2000; 2000CA-2307010.

(INFE-) INFECTIO DIAGNOSTIC (IDI) INC.

Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;

Picard FJ, Roy PH;

WPI; 2001-245006/25.

Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitica species in a test sample -
XX
PS Claim 27; Page 961; 1580pp; English.

The present invention describes a method for generating a repository of
CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitica
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archaeal, bacterial, fungal

CC and parasitica species, genus, family and group. A nucleic acid (I)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexa nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (1) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonas group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.

SQ Sequence 1212 BP; 391 A; 269 C; 232 G; 320 T; 0 other;

Query Match 6.0%; Score 120; DB 22; Length 1212;
Best Local Similarity 100.0%; Pred. No. 2e-51;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1842 CTGAAAGTTCAAGCTCATCATGTTCACTTCAAGTCTAGCTCAACCACTCCAA 1901
|
DB 1093 ctgaaagttcaagctcatcatcagatagttcaacttcacagctcagctcaaccactccaa 1152

QY 1902 GCACAATTAATAGTACGACGTACCAATCTTAACATAATAGCAGCAATATCAATCAACCC 1961
|
DB 1153 gcacaataataagtagcagactaccatcctaataataagcaacaatacaatacaacc 1212

RESULT 42

AAH02064
ID AAH02064 standard; DNA; 1242 BP.

AAH02064;

24-JUL-2001 (first entry)

Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2057.

Species specific; genus specific; family specific; probe: detection;
KW identification: algal; archaeal; bacterial; fungal; parasitica;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.

Streptococcus pneumoniae.

WO200123604-A2.

05-APR-2001.

28-SEP-2000; 2000WO-CA01150.

28-SEP-1999; 99CA-2283458.

19-MAY-2000; 2000CA-2307010.

(INFE-) INFECTIO DIAGNOSTIC (IDI) INC.

Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;

Picard FJ, Roy PH;

WPI; 2001-245006/25.

Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitica species in a test sample -
XX
PS Claim 27; Page 1471; 1580pp; English.


```
XX XX Streptococcus pneumoniae.
XX OS
XX PN W0200123604-A2.
XX XX
XX PD 05-APR-2001.
XX XX
XX PF 28-SEP-2000; 2000MO-CA01150.
XX XX
XX PR 28-SEP-1999; 99CA-2283458.
XX PR 19-MAR-2000; 2000CA-2307010.
XX XX
XX PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX XX
XX PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
XX PI Picard FJ, Roy PH;
XX DR WPI: 2001-245006/25.
XX XX
XX PT Nucleic acid sequences are used to generate universal probes and
XX PT primers which can be used to identify and detect the presence of algal,
XX PT archaeal, bacterial, fungal and parasitica species in a test sample -
XX XX
XX PS Claim 27: Page 1472; 1580pp; English.
XX XX
XX CC The present invention describes a method for generating a repertory of
XX CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
XX CC and/or primers are derived. The method comprises amplifying the nucleic
XX CC acids of determined algal, archaeal, bacterial, fungal and parasitica
XX CC species with a combination of defined primer pairs. The method can be
XX CC used for producing probes and/or primers for detecting one or more
XX CC related microorganisms e.g. algae, archaea, bacteria, fungi and
XX CC parasites, for universal detection and for specific and ubiquitous
XX CC detection and identification of an algal, archaeal, bacterial, fungal
XX CC and parasitica species, genus, family and group. A nucleic acid (I)
XX CC obtained using the method of the invention can be used for the universal
XX CC detection of any bacterium, fungus or parasite in a sample and for the
XX CC detection of at least one antimicrobial agent resistance gene or at
XX CC least one toxin gene. hexa nucleic acids are used for the specific and
XX CC ubiquitous detection and for identification of Streptococcus pneumoniae.
XX CC (I) can be used to design a therapeutic agent which is effective against
XX CC microorganisms. Microbial species or genus or family or phylum or group
XX CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
XX CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
XX CC Mycobacteriaceae family, Pseudomonas group, Streptococcus sp.,
XX CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
XX CC provides faster results than substrate specificity tests as results can
XX CC be determined in an hour and improved accuracy is also achieved.
XX CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
XX CC which are given in the exemplification of the present invention.
XX XX
XX SQ Sequence 1225 BP; 378 A; 278 C; 248 G; 321 T; 0 other;
XX XX
XX Query Match 5.2%; Score 104; DB 22; Length 1225;
XX Best Local Similarity 100.0%; Pred. No. 3.7e-43;
XX Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX QY 755 GAACGCTGAAAACATCTGCTGGAGATATTACATACAGACGATACGCTTACACACAC 814
XX |
XX DB 1 gaagctcaaaaacatctgctggatattcaacaatacagacgatacgttgcctaccagac 60
XX |
XX QY 815 GATGAATTTGCAAGTCGCTTACCATTTGATGATGTTTCTTACACG 858
XX |
XX DB 61 gaaagattgcagtcgcttaccatgttgatttcttaacgg 104
XX |
XX RESULT 45
XX ID AA235940 standard; DNA: 1260 BP.
XX AC AA235940;
XX XX
```

```
DT 07-FEB-2000 (first entry)
XX XX
XX DE Streptococcus pneumoniae pbp1a TER isolate b) nucleotide sequence.
XX XX
XX KW Streptococcus pneumoniae; penicillin binding protein; pbp2b; pbp1a;
XX KW transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
XX KW detection; identification; pneumococcal meningitis; ss.
XX XX
XX OS Streptococcus pneumoniae.
XX XX
XX PN ZA9807024-A.
XX XX
XX PD 28-APR-1999.
XX XX
XX PF 05-AUG-1998; 98ZA-0007024.
XX XX
XX PR 01-AUG-1997; 97ZA-0006886.
XX XX
XX PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
XX PA (UYMI-) UNIV WITWATERSRAND.
XX PA (MEDI-) MEDICAL RES COUNCIL.
XX PI Klugman KP, Smith AM, Du Plessis M;
XX PI WPI: 1999-601770/51.
XX DR WPI: 1999-601770/51.
XX XX
XX PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
XX PT useful for the diagnosis of pneumococcal meningitis
XX XX
XX PS Claim 11; Fig 4; 63pp; English.
XX XX
XX CC A polymerase chain reaction (PCR) assays have been developed for
XX CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
XX CC using primers based on the penicillin binding protein 2b (pbp2b) gene
XX CC and the pbp1a gene. The products and methods can be used for detecting
XX CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
XX CC used for simultaneously diagnosing pneumococcal meningitis and
XX CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
XX CC The methods can be used for detecting S. pneumoniae strains resistant
XX CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
XX CC The assays can be adapted to detect other pathogens causing meningitis.
XX CC The assays can be used to detect an antibiotic resistant strain of
XX CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
XX CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
XX CC a 224 bp product. The present sequence represents a Streptococcus
XX CC pneumoniae pbp1a transpeptidase encoding region (TER) isolate nucleotide
XX CC sequence from the present invention.
XX XX
XX SQ Sequence 1260 BP; 394 A; 264 C; 258 G; 344 T; 0 other;
XX XX
XX Query Match 5.1%; Score 101; DB 20; Length 1260;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-41;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX QY 575 GAGACGATGGAAGACGATCAATACCAATTAATGATGATGATCAAGTCAATCA 634
XX |
XX DB 1 gaacagataggaagaagcgaatcaacacatctcgtatgagcacaagaagctcaaatca 60
XX |
XX QY 635 GCAGTAATTAACCTGCTTACATGATGATTAATTAATCAAGCA 675
XX |
XX DB 61 gcaagtaattaccctcgtctacatgataattaccctaagga 101
XX |
XX RESULT 46
XX ID AA235941
XX XX
XX AC AA235941;
XX AC 07-FEB-2000 (first entry)
XX DE Streptococcus pneumoniae pbp1a TER isolate c) nucleotide sequence.
XX XX
```


XX Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;
KW Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;
KM translation elongation factor G; RecA recombinase; resistance;
XX catalytic subunit of proton-translocating ATPase; antimicrobial;
XX vaccine; primer: ds.
OS Streptococcus pneumoniae.
XX WO200123604-A2.
PN 28-SEP-2000; 2000WO-CA01150.
XX 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
PI WPI: 2001-245006/25.
DR Nucleic acid sequences are used to generate universal probes and
XX primers which can be used to identify and detect the presence of algal,
XX archaeal, bacterial, fungal and parasitological species in a test sample -
XX Disclosure: Page 1466; 1580pp; English.
PS The present invention describes a method for generating a repository of
XX nucleic acids of tuf, fus, atpD and/or recA genes from which probes
XX and/or primers are derived. The method comprises amplifying the nucleic
XX acids of determined algal, archaeal, bacterial, fungal and parasitological
XX species with a combination of defined primer pairs. The method can be
XX used for producing probes and/or primers for detecting one or more
XX related microorganisms e.g. algae, archaea, bacteria, fungi and
XX parasites, for universal detection and for specific and ubiquitous
XX detection and identification of an algal, archaeal, bacterial, fungal
XX and parasitological species, genus, family and group. A nucleic acid (I)
XX obtained using the method of the invention can be used for the universal
XX detection of any bacterium, fungus or parasite in a sample and for the
XX detection of at least one antimicrobial agent resistance gene or at
XX least one toxin gene. hexA nucleic acids are used for the specific and
XX ubiquitous detection and for identification of Streptococcus pneumoniae.
XX (I) can be used to design a therapeutic agent which is effective against
XX microorganisms. Microbial species or genus or family or phylum or group
XX which can be detected include Abiotrophia adiacens, Bordetella sp.,
XX Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
XX Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
XX Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
XX can be determined in an hour and improved accuracy is also achieved.
XX AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
XX are given in the exemplification of the present invention.
SQ Sequence 930 BP; 282 A; 195 C; 192 G; 261 T; 0 other;
Query Match 4.9%; Score 97; DB 22; Length 930;
Best Local Similarity 100.0%; Pred. No. 1.5e-39;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1377 AAAAAGTTCCTATATGTCGACGTCGTCCTATGAAGCAAGCAAGCTATATGATCA 1436
DB 608 aaaaagatctcctaatagtcgcgactcgtgcatagaagaacagacctataatgatga 667
OY 1437 CCGACATGATGAAAACAGCTTGACTTATGAACTGG 1473
DB 668 ccgacatgatgaacacagctctgactatggaactg 704
RESULT 48
AAH02057
ID AAH02057 standard; DNA; 930 BP.
XX
XX AAH02057;
XX
XX 24-JUL-2001 (first entry)
XX
XX Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2050.
XX
XX Species specific; genus specific; family specific; probe; detection;
KW identification; algal; archaeal; bacterial; fungal; parasitological;

XX microorganism; diagnosis; translation elongation factor Tuf; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
XX catalytic subunit of proton-translocating ATPase; antimicrobial;
XX vaccine; primer: ds.
OS Streptococcus pneumoniae.
XX WO200123604-A2.
PN 28-SEP-2000; 2000WO-CA01150.
XX 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
PI WPI: 2001-245006/25.
DR Nucleic acid sequences are used to generate universal probes and
XX primers which can be used to identify and detect the presence of algal,
XX archaeal, bacterial, fungal and parasitological species in a test sample -
XX Disclosure: Page 1466; 1580pp; English.
PS The present invention describes a method for generating a repository of
XX nucleic acids of tuf, fus, atpD and/or recA genes from which probes
XX and/or primers are derived. The method comprises amplifying the nucleic
XX acids of determined algal, archaeal, bacterial, fungal and parasitological
XX species with a combination of defined primer pairs. The method can be
XX used for producing probes and/or primers for detecting one or more
XX related microorganisms e.g. algae, archaea, bacteria, fungi and
XX parasites, for universal detection and for specific and ubiquitous
XX detection and identification of an algal, archaeal, bacterial, fungal
XX and parasitological species, genus, family and group. A nucleic acid (I)
XX obtained using the method of the invention can be used for the universal
XX detection of any bacterium, fungus or parasite in a sample and for the
XX detection of at least one antimicrobial agent resistance gene or at
XX least one toxin gene. hexA nucleic acids are used for the specific and
XX ubiquitous detection and for identification of Streptococcus pneumoniae.
XX (I) can be used to design a therapeutic agent which is effective against
XX microorganisms. Microbial species or genus or family or phylum or group
XX which can be detected include Abiotrophia adiacens, Bordetella sp.,
XX Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
XX Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
XX Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
XX can be determined in an hour and improved accuracy is also achieved.
XX AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
XX are given in the exemplification of the present invention.
SQ Sequence 930 BP; 282 A; 195 C; 192 G; 261 T; 0 other;
Query Match 4.9%; Score 97; DB 22; Length 930;
Best Local Similarity 100.0%; Pred. No. 1.5e-39;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1377 AAAAAGTTCCTATATGTCGACGTCGTCCTATGAAGCAAGCAAGCTATATGATCA 1436
DB 608 aaaaagatctcctaatagtcgcgactcgtgcatagaagaacagacctataatgatga 667
OY 1437 CCGACATGATGAAAACAGCTTGACTTATGAACTGG 1473
DB 668 ccgacatgatgaacacagctctgactatggaactg 704
RESULT 48
AAH02057
ID AAH02057 standard; DNA; 930 BP.
XX
XX AAH02057;
XX
XX 24-JUL-2001 (first entry)
XX
XX Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2050.
XX
XX Species specific; genus specific; family specific; probe; detection;
KW identification; algal; archaeal; bacterial; fungal; parasitological;

ID AA235942 standard; DNA: 1260 BP.
XX
AC AA235942;
XX
DT 07-FEB-2000 (first entry)
XX
DE Streptococcus pneumoniae pbp1a TER isolate d) nucleotide sequence.
XX
DE Streptococcus pneumoniae: penicillin binding protein: pbp2B; pbp1a;
XX transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
KM detection; identification; pneumococcal meningitis; ss.
XX
OS Streptococcus pneumoniae.
XX
PN ZA9807024-A.
XX
PD 28-APR-1999.
XX
PF 05-AUG-1998; 982A-0007024.
XX
PR 01-AUG-1997; 972A-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UWI-) UNIV WITWATERSRAND.
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smilt AM, Du Plessis M;
XX WPI; 1999-601770/51.
DR
XX
PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis
XX
XX
PS Claim 11; Fig 4; 63pp; English.
XX
XX A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1a gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1a transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX
SO Sequence 1260 BP; 404 A; 266 C; 254 G; 336 T; 0 other;

Query Match 4.9%; Score 97; DB 20; Length 1260;
Best Local Similarity 100.0%; Pred. No. 1.5e-39;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1377 AAAAAGCTTCTATGTGGAAGTGTGCATGAAGAAGACAGCCTATATGATGA 1436
DB 803 aaaaagagttcttaagtgcgaactcgtgcataagaagaacagacactatagatga 862
OY 1437 CCGACATGATGAAGACGCTGTGACTATANGAAGCTGG 1473
DB 863 ccgacatgataaagaacgctctgactatggaactgg 899

RESULT 49
AAH02071
ID AAH02071 standard; DNA: 782 BP.
XX
AC AAH02071;

XX
DE 24-JUL-2001 (first entry)
XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2064.
XX
XX
XX Species specific; genus specific; family specific; probe; detection;
KW identification; algal; archaeal; bacterial; fungal; parasitical;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
XX vaccine; primer; ds.
XX
OS Streptococcus pneumoniae.
XX
PN W0200123604-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-CA01150.
XX
PR 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX
PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
XX WPI; 2001-245006/25.
DR
XX
XX
XX Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitical species in a test sample -
XX
XX
PS Claim 27; Page 1476; 1580pp; English.
XX
XX The present invention describes a method for generating a repository of
CC nucleic acids of ruf, fus, atpd and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitical
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archaeal, bacterial, fungal
CC and parasitical species, genus, family and group. A nucleic acid (II)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexa nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (I) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Adenoviridae, Bacterioides, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
SO Sequence 782 BP; 247 A; 181 C; 156 G; 198 T; 0 other;

Query Match 4.5%; Score 90; DB 22; Length 782;
Best Local Similarity 100.0%; Pred. No. 6.2e-36;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 965 ATCAGACATATGCTCTGCTTGGAGTACGAGTGTCTACAGTTAACTGCTACTATGCTT 1024
DB 163 atcacagactatgctctgcttgagtagtactgctcagcatcactgctactatcgtt 222

QY	1025	CACGATGAGCCCTATTA	CTACCC	TGGGACA	1054
Db	223	cacgatgagccctata	ctactac	cctg	252

RESULT 50

AAH01655
ID AAH01655 standard; DNA; 813 BP.

AC AAH01655;

DT 24-JUL-2001 (first entry)

Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1648.

KM Species specific; genus specific; family specific; probe; detection;
KM Identification; algal; archaeal; bacterial; fungal; parasitical;
KM Microorganism; diagnosis; translation elongation factor Tu; toxin;
KM translation elongation factor G; RecA recombinase; resistance;
KM catalytic subunit of proton-translocating ATPase; antimicrobial;
KM vaccine; primer; ds.

Streptococcus pneumoniae.

PN WO200123604-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-CA01150.

PR 28-SEP-1999; 99CA-2283458 -
PR 19-MAY-2000; 2000CA-3307010 -

XX
XX
(INFE-) INFECTION DIAGNOSTICPA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX

PI Bergeron MG, Bolssinot M, Huletsky A, Menard C, Ouellette M;
PI Picard ET, Roy PH:

DR WPI: 2001-245006/25.

PT Nucleic acid sequences are used to generate universal probes and
 PT Primers which can be used to identify and detect the presence of algal
 PT archaeal, bacterial, fungal and parasitical species in a test sample -
 PS Claim 27; Page 1255; 1560pp: English.
 XX

The present invention describes a method for generating a repository of nucleic acids of tuf, fus, atpd and/or recA genes from which probes and/or primers are derived. The method comprises amplifying the nucleic acids of determined algal, archaeal, bacterial, fungal and parasitical species with a combination of defined primer pairs. The method can be used for producing probes and/or primers for detecting one or more related microorganisms e.g. algae, archaea, bacteria, fungi and parasites, for universal detection and for specific and ubiquitous detection and identification of an algal, archaeal, bacterial, fungal and parasitical species, genus, family and group. A nucleic acid (i) obtained using the method of the invention can be used for the universal detection of any bacterium, fungus or parasite in a sample and for the detection of at least one antimicrobial agent resistance gene or at least one toxin gene. hexa nucleic acids are used for the specific and ubiquitous detection and for identification of *Streptococcus pneumoniae*. (i) can be used to design a therapeutic agent which is effective against microorganisms. Microbial species or genus or phylum or group which can be detected include *Blattaria adiacens*, *Bordetella* sp., *Corynebacterium* sp., *Enterobacteriaceae* group, *Escherichia coli*, *Mycobacteriaceae* family, *Pseudomonas* group, *Streptococcus* sp., *Neisseria gonorrhoeae* and *Staphylococcus* sp.. using DNA based tests can provides faster results than substrate specificity tests as results can be determined in an hour and improved accuracy is also achieved. AAH00010 to AAH02304 represent nucleotide sequences and primers/probes which are given in the exemplification of the present invention.

SQ Sequence 813 BP; 253 A; 186 C; 164 G; 210 T; 0 other;

Query Match	4.5%;	Score 90;	DB 22;	Length 813;
Best Local Similarity	100.0%;	Pred. No. 6.2e+36;		
Matches 90;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Oy 965 ATACAGACTATGCTCCTCGCTTGAGTACGGTGTCTACGATTCAACTGCTACTATCGTT 1024
|||
Db 169 atcacagactatgctcctcgcttgagtagtgygtctacgaattcaactgcgtactatcggt 228

QY	1025	CACGATGAGCCCTATACTACCCCTGGGACA	1054
Db	229	cacgatgagccctatactaacctgagaca	258

```
Search completed: June 13, 2002, 12:42:27
Job time: 6595 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 08:39:33 : Search time 2129.67 Seconds
(without alignments)
12668.823 Million cell updates/sec

Title: US-08-961-083-1

Perfect score: 1999

Sequence: 1 TAAATCTACGACATATAA.....ATCCTCAGCAGCAACCA 1999

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 15

Total number of hits satisfying chosen parameters: 25133

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	1.2	928	10 BG334672	BG334672 602460969
2	22	1.1	553	12 AO701777	AO701777 HS_2134_A
3	22	1.1	909	12 CNS0363A	AL229519 Tetracodon
4	21	1.1	389	9 AL632585	AL632585
5	21	1.1	422	9 AI397462	AI397462 fbl1808.x
6	21	1.1	704	9 AL629478	AL629478
7	21	1.1	1025	12 CNS06N80	AL629478 AL629478
8	21	1.1	1078	12 CNS06NM7	AL629478 T3 end of
9	20	1.0	237	10 BF174868	AL406949 T7 end of
10	20	1.0	287	9 AA317576	BF174868 MYE3951 M
11	20	1.0	295	9 BB551207	AA317576 EST19524
12	20	1.0	360	10 BF016990	BB551207 BB551207
13	20	1.0	367	10 BF016981	BF016990 ux19d08.y
14	20	1.0	402	10 BM173589	BF016981 ux19c08.y
15	20	1.0	405	10 BF016982	BM173589 900943 Av
16	20	1.0	406	12 AO445724	BF016982 ux19c09.y
17	20	1.0	440	10 BG017796	AO445724 nbxb0054J

18	20	1.0	442	9	BB824552	BB824552
19	20	1.0	443	10	BE566388	BE566388 601339984
20	20	1.0	446	10	BG942702	BG942702 ax28e04.x
21	20	1.0	450	9	BB750480	BB750480 BB750480
22	20	1.0	493	9	AA280963	AA280963 zs97q12.x
23	20	1.0	526	9	AW621116	AW621116 ur53h07.y
24	20	1.0	526	12	AO710693	AO710693 HS_5344_A
25	20	1.0	535	12	AZ824983	AZ824983 2M0099J24
26	20	1.0	563	10	BG723460	BG723460 602694145
27	20	1.0	587	10	BG503218	BG503218 602651318
28	20	1.0	588	10	BG539230	BG539230 60267826
29	20	1.0	642	9	AA606845	AA606845 ym94h07.r
30	20	1.0	650	10	BG492386	BG492386 602536313
31	20	1.0	661	9	BB498524	BB498524 BB498524
32	20	1.0	669	10	BF212414	BF212414 601813376
33	20	1.0	677	10	BG98708	BG98708 HOA56-1-G
34	20	1.0	705	9	AA152866	AA152866 mq56c03.r
35	20	1.0	710	10	BF658632	BF658632 maa98h01.
36	20	1.0	712	12	AG099717	AG099717 Pan trogl
37	20	1.0	731	10	BG506942	BG506942 601861569
38	20	1.0	737	12	AZ354492	AZ354492 1M0093715
39	20	1.0	787	10	BF239667	BF239667 601906930
40	20	1.0	788	10	BE889059	BE889059 601513993
41	20	1.0	803	10	BI690754	BI690754 603314371
42	20	1.0	841	10	BF667800	BF667800 602122620
43	20	1.0	861	10	BE383813	BE383813 601298315
44	20	1.0	873	12	BH466047	BH466047 BCGV80TR
45	20	1.0	888	10	BF670847	BF670847 602149929
46	20	1.0	896	10	BF696798	BF696798 602125331
47	20	1.0	904	10	BF247101	BF247101 601854544
48	20	1.0	917	10	BF697265	BF697265 602129793
49	20	1.0	938	10	BG282996	BG282996 602406185
50	20	1.0	969	12	AO688268	AO688268 nbxb0077M
51	20	1.0	1021	10	BE964433	BE964433 601658140
52	20	1.0	1101	10	BM480145	BM480145 AGENCOURT
53	20	1.0	1151	10	BF797518	BF797518 602257187
54	20	1.0	1537	10	BM457339	BM457339 AGENCOURT
55	20	1.0	3519	11	AK010591	AK010591 Mus muscu
56	19	1.0	120	12	TA129D01P	TA129D01P
57	19	1.0	225	9	AV377407	AV377407
58	19	1.0	312	9	AV299242	AV299242
59	19	1.0	319	10	BG801896	BG801896 0122-44 M
60	19	1.0	325	12	AO567904	AO567904 HS_2117_A
61	19	1.0	367	9	AM61079	AM61079 PM4-CT029
62	19	1.0	375	10	BM264807	BM264807 fw62g01.x
63	19	1.0	401	10	Z29739	Z29739 ARTS2183 Or
64	19	1.0	407	10	BE925196	BE925196 CM2-AN007
65	19	1.0	408	9	BE039091	BE039091 AB09E06 A
66	19	1.0	411	12	BB6567	BB6567 PC111-24C1
67	19	1.0	433	10	BG077137	BG077137 H3011C04-
68	19	1.0	450	12	AZ143510	AZ143510 SP_0018_A
69	19	1.0	465	12	AO013306	AO013306 PC111-24
70	19	1.0	470	12	AZ086245	AZ086245 RPCI-23-4
71	19	1.0	484	10	BF994356	BF994356 MR0-GN017
72	19	1.0	489	10	BF775254	BF775254 285355 MA
73	19	1.0	497	12	AZ763586	AZ763586 1M0559B08
74	19	1.0	501	12	BH217548	BH217548 1006056C0
75	19	1.0	527	12	AZ759820	AZ759820 1M0552L24
76	19	1.0	531	10	BM334118	BM334118 MEST13-H
77	19	1.0	541	10	BI974963	BI974963 sal15c11
78	19	1.0	543	10	BE460169	BE460169 EST115461
79	19	1.0	566	10	BT705962	BT705962 f98F03.y
80	19	1.0	581	10	BE459685	BE459685 EST14977
81	19	1.0	581	12	AO788778	AO788778 HS_3187_B
82	19	1.0	593	12	BE458253	BE458253 EST13593
83	19	1.0	608	12	BH259200	BH259200 CH230-29B
84	19	1.0	613	12	AO385601	AO385601 PC111-13
85	19	1.0	626	10	BM490620	BM490620 P922n.p0
86	19	1.0	635	12	AZ409586	AZ409586 1M0181615
87	19	1.0	660	12	BH423208	BH423208 BCGMFI2TF
88	19	1.0	665	12	BH062477	BH062477 RPCI-24-3
89	19	1.0	677	9	AI588624	AI588624 fb97d10.y
90	19	1.0	677	10	BI661446	BI661446 603304741

91	19	1.0	681	10	BI099954	BI099954 602885082	C 164	18	0.9	403	12	AQ253985	AQ253985 HS-2225-A
92	19	1.0	698	12	AZ908847	PCIC-24-1	C 165	18	0.9	406	9	AM168430	AM168430 x183d04.x
93	19	1.0	721	10	BF581163	602100340	C 166	18	0.9	406	9	BM303196	BM303196 f148d09.x
94	19	1.0	725	12	AZ629815	1M0483J06	C 167	18	0.9	407	10	BI578309	BI578309 RE72162.5
95	19	1.0	752	12	BH244978	AB28A71TR	C 168	18	0.9	408	10	BI485888	BI485888 RE69241.5
96	19	1.0	758	10	BG963476	602831137	C 169	18	0.9	411	9	AV729734	AV729734 AV729734
97	19	1.0	776	10	BI102190	602887975	C 170	18	0.9	421	10	BI484107	BI484107 RE67141.5
98	19	1.0	785	12	BH583112	B06EN26TF	C 171	18	0.9	421	9	AU208871	AU208871 AU208871
99	19	1.0	801	12	BH482780	BOHFE96TF	C 172	18	0.9	423	10	BI67181	BI67181 CRK-HSP-204
100	19	1.0	805	12	BH511534	B06BD36TR	C 173	18	0.9	423	12	BH0034	BH0034 CIT-HSP-204
101	19	1.0	809	12	BP236718	BF262718 602028013	C 174	18	0.9	427	10	BI005041	BI005041 PM3-HN007
102	19	1.0	859	12	BH111314	RPCT-24-3	C 175	18	0.9	430	10	AM477317	AM477317 ga45c11.y
103	19	1.0	876	10	BF263156	HY-CEA000	C 176	18	0.9	431	12	AQ584924	AQ584924 RPCT-11-4
104	19	1.0	906	12	AZ689698	ENTK085TF	C 177	18	0.9	436	10	C82907	C82907 C82907 rabd
105	19	1.0	910	12	AZ550500	ENTREC19TF	C 178	18	0.9	436	10	C83763	C83763 C83763 rabd
106	19	1.0	924	10	BF584742	602908492	C 179	18	0.9	437	10	W35196	W35196 c7d0d06.r1
107	19	1.0	940	10	BG98341	602397127	C 180	18	0.9	438	9	AI475337	AI475337 f182a04.x
108	19	1.0	1038	12	CNS01HM	AL14659 Anopheles	C 181	18	0.9	438	10	BM440956	BM440956 EBED02-50
109	19	1.0	2199	11	AK019222	AK019222 Mus muscu	C 182	18	0.9	438	10	R73286	R73286 yj92d01.r1
110	18	0.9	117	10	T63516	T63516 yc07a03.s1	C 183	18	0.9	438	12	AZ475157	AZ475157 IM0293B11
111	18	0.9	124	12	BH417338	1007054D0	C 184	18	0.9	439	10	C06769	C06769 C06769 Rat
112	18	0.9	174	12	AO269185	RPCT11-75	C 185	18	0.9	441	12	AO842111	AO842111 T134A07.S
113	18	0.9	183	10	BG968819	602836284	C 186	18	0.9	442	10	BF222120	BF222120 T942f03.x
114	18	0.9	187	9	AA711964	AA711964 v029c04.r	C 187	18	0.9	442	12	AQ265185	AQ265185 CTBI-EL-
115	18	0.9	207	12	B96425	F19B14TR IG	C 188	18	0.9	449	12	AO695311	AO695311 HS-2142-A
116	18	0.9	208	12	AM859435	AM859435 MRI-CT035	C 189	18	0.9	450	10	BE497833	BE497833 WHE0957.F
117	18	0.9	208	12	BH206244	SM1-49H7.	C 190	18	0.9	452	9	AA293412	AA293412 c153f09.r
118	18	0.9	217	9	AV334982	AV334982	C 191	18	0.9	454	9	AM468215	AM468215 hc34c01.x
119	18	0.9	234	9	AM859384	AM859384 MRI-CT035	C 192	18	0.9	459	9	AW770104	AW770104 h183g01.x
120	18	0.9	253	9	AA356650	EST65492	C 193	18	0.9	460	10	W95864	W95864 ze08a01.s1
121	18	0.9	257	9	AA322319	EST24941	C 194	18	0.9	461	10	BI797071	BI797071 H063H02.E
122	18	0.9	271	10	NS59046	NS59046 yv60c12.r1	C 195	18	0.9	461	12	AZ811979	AZ811979 J2M0087D14
123	18	0.9	272	10	BG940864	ax07h09.y	C 196	18	0.9	462	9	AI588336	AI588336 fb99e05.x
124	18	0.9	279	9	AA934680	0080b12.s	C 197	18	0.9	463	10	BE411391	BE411391 925005C08
125	18	0.9	283	12	AO068440	HS-2243.B	C 198	18	0.9	464	10	BI164963	BI164963 RE04744.5
126	18	0.9	283	10	BE938266	BE938266 fm74e06.y	C 199	18	0.9	469	9	AI811634	AI811634 tw44a10.x
127	18	0.9	286	10	BI005131	PM3-HN007	C 200	18	0.9	469	9	AI840859	AI840859 UI-M-AHO-
128	18	0.9	288	10	BB228355	BB228355	C 201	18	0.9	473	10	BI137007	BI137007 F078P35Y
129	18	0.9	292	9	BE074170	RC4-ET056	C 202	18	0.9	473	10	BI124998	BI124998 BI124998
130	18	0.9	302	9	AI945138	bs09d08.y	C 203	18	0.9	474	10	BE442227	BE442227 925014H07
131	18	0.9	311	10	BM164426	BM164426	C 204	18	0.9	475	9	AI944429	AI944429 bs01c09.y
132	18	0.9	311	10	BM238980	BM238980 K0526G08-	C 205	18	0.9	477	10	BG639388	BG639388 LP07744.3
133	18	0.9	323	12	AQ382777	RPCT11-13	C 206	18	0.9	482	10	R95815	R95815 yq50e07.r1
134	18	0.9	325	9	AM565951	s18bh03.y	C 207	18	0.9	483	10	BI592216	BI592216 RH09855.5
135	18	0.9	325	10	BG940862	ax07h09.y	C 208	18	0.9	486	9	AA254066	AA254066 va10f05.r
136	18	0.9	332	9	AV689732	AV689732	C 209	18	0.9	486	9	AA589206	AA589206 VK24e01.r
137	18	0.9	334	10	BI473730	BI473730 fP41f04.y	C 210	18	0.9	486	9	AA589206	AA589206 VK24e01.r
138	18	0.9	347	9	AA658202	AA658202 nu20D06.S	C 211	18	0.9	487	9	BM239585	BM239585 K0537A12-
139	18	0.9	355	12	AZ904101	RPCT-24-1	C 212	18	0.9	487	9	AA395358	AA395358 27155.Lam
140	18	0.9	358	9	AM258934	AM258934 um75a08.y	C 213	18	0.9	490	10	BF623829	BF623829 HVSMEa000
141	18	0.9	360	9	AI301700	AI301700 qm36q08.x	C 214	18	0.9	491	9	AI654525	AI654525 tg91g06.x
142	18	0.9	367	10	BF431122	BF431122 260495.MA	C 215	18	0.9	491	10	BE857967	BE857967 7673d04.x
143	18	0.9	368	9	AM273218	AM273218 x335e08.x	C 216	18	0.9	494	9	AA576752	AA576752 nm81g08.s
144	18	0.9	369	10	BI228681	BI228681 RE26305.5	C 217	18	0.9	496	10	BG849419	BG849419 1024025C0
145	18	0.9	369	12	AZ503052	AZ503052 IM0342111	C 218	18	0.9	496	10	BI366957	BI366957 RE52612.5
146	18	0.9	370	9	AM275317	AM275317 xv80h12.x	C 219	18	0.9	497	12	AO941496	AO941496 Sheared.D
147	18	0.9	372	12	BH232336	BH232336 100616780	C 220	18	0.9	500	10	BI482175	BI482175 BE64833.5
148	18	0.9	374	9	AA504438	AA504438 aa59g11.r	C 221	18	0.9	500	12	AZ021229	AZ021229 RPCT-23-3
149	18	0.9	376	10	C69781	C69781 C69781 YUJ1	C 222	18	0.9	504	9	AM198586	AM198586 p115.pK0
150	18	0.9	377	9	AV729733	AV729733	C 223	18	0.9	505	9	AI742888	AI742888 wq47e02.x
151	18	0.9	383	9	AV188421	AV188421	C 224	18	0.9	505	10	BI239057	BI239057 RE35447.5
152	18	0.9	383	12	AZ904732	AZ904732 RPCT-24-1	C 225	18	0.9	507	12	BF654492	BF654492 278653.MA
153	18	0.9	384	9	AA164846	AA164846 zp02b11.f	C 226	18	0.9	508	12	AZ846966	AZ846966 2M0147C23
154	18	0.9	390	12	AA536345	AA536345 LD16910.5	C 227	18	0.9	510	12	AO700726	AO700726 HS-5321.A
155	18	0.9	393	9	AA536345	AA536345 LD16910.5	C 228	18	0.9	510	10	BI235507	BI235507 RE3159.5
156	18	0.9	393	10	BE834619	BE834619 MRI-FN000	C 229	18	0.9	511	10	BI136700	BI136700 F072P19Y
157	18	0.9	396	12	AO511746	AO511746 HS-5119.A	C 230	18	0.9	512	9	AI953689	AI953689 wq47c01.x
158	18	0.9	399	12	AO152510	AO152510 HS-3110.B	C 231	18	0.9	513	10	BI367740	BI367740 RE53572.5
159	18	0.9	400	9	AI394888	AI394888 MA001297.	C 232	18	0.9	517	10	BI004814	BI004814 BT004814
160	18	0.9	401	9	AI385055	AI385055 MA002302.	C 233	18	0.9	517	10	BI112673	BI112673 BT112673
161	18	0.9	403	9	AA025915	AA025915 ze91c03.r	C 234	18	0.9	518	10	BI169087	BI169087 RE10064.5
162	18	0.9	403	10	Z34597	Z34597 ATTS3357.Ve	C 235	18	0.9	518	10	BI804887	BI804887 S001F07.S
163	18	0.9	403	10	Z34597	Z34597 ATTS3357.Ve	C 236	18	0.9	518	10	BM301346	BM301346 MCR041F08

C 237	18	0.9	520	12	A2172702	SP_0122-B	310	18	0.9	615	9	A0034816	A0034816
C 238	18	0.9	521	10	B1368754	RE54779.5	311	18	0.9	616	12	A2947246	A2947246
C 239	18	0.9	524	12	A0017893	CIT-HSP-2	312	18	0.9	617	10	B1212951	B1212951
C 240	18	0.9	525	10	BE672082	7a51c01.x	313	18	0.9	617	12	A2800372	A2800372
C 241	18	0.9	527	12	BE4374	CIT-HSP-202	314	18	0.9	618	9	AV401322	AV401322
C 242	18	0.9	529	10	B1227672	RE24781.5	315	18	0.9	618	9	BD498288	BD498288
C 243	18	0.9	529	10	BM116635	BM116635 L0839B04-	316	18	0.9	619	10	B1364444	B1364444
C 244	18	0.9	530	12	AM055074	AM055074 WY9C609.x	317	18	0.9	619	10	B1216624	B1216624
C 245	18	0.9	531	10	BG150430	BG150430 7k01a02.x	318	18	0.9	619	10	B1374865	B1374865
C 246	18	0.9	532	9	AI587545	AI587545 tF52b07.x	319	18	0.9	619	12	A2334737	A2334737
C 247	18	0.9	532	9	AI587545	AI587545 tF52b07.x	320	18	0.9	620	10	B1171092	B1171092
C 248	18	0.9	532	12	AE501525	AE501525 1M03401122	321	18	0.9	621	10	B1364570	B1364570
C 249	18	0.9	534	12	BE590370	BE590370 nab23g01.	322	18	0.9	622	10	AL577457	AL577457
C 250	18	0.9	534	12	BE5229	BE5229 RPTC11-30M7	323	18	0.9	622	12	AQ449509	AQ449509
C 251	18	0.9	535	9	AI483218	AI483218 EST242695	324	18	0.9	623	10	BM159288	BM159288
C 252	18	0.9	535	9	AI638520	AI638520 tF08f03.x	325	18	0.9	624	9	BB630750	BB630750
C 253	18	0.9	537	9	AI608110	AI608110 vA10f05.y	326	18	0.9	627	10	B1162430	B1162430
C 254	18	0.9	540	10	B1375957	B1375957 RE63866.5	327	18	0.9	629	10	BF863372	BF863372
C 255	18	0.9	540	12	AQ753968	HS_5398_A	328	18	0.9	630	12	BH056545	BH056545
C 256	18	0.9	542	10	BJ116355	BJ116355 BJ116355	329	18	0.9	631	9	AI357593	AI357593
C 257	18	0.9	542	10	BF196828	BF196828 7n07d04.x	330	18	0.9	631	10	BJ006434	BJ006434
C 258	18	0.9	545	10	BE429899	BE429899 TAS004.H1	331	18	0.9	632	9	AI774507	AI774507
C 259	18	0.9	547	9	AI675148	AI675148 wC09h11.x	332	18	0.9	633	9	AA201744	AA201744
C 260	18	0.9	547	10	BM407754	BM407754 EST582081	333	18	0.9	634	9	AM696903	AM696903
C 261	18	0.9	550	12	AQ281037	AQ281037 RPTC11-81	334	18	0.9	634	9	BB208421	BB208421
C 262	18	0.9	551	10	BI486090	BI486090 RE69463.5	335	18	0.9	639	10	B1214545	B1214545
C 263	18	0.9	551	10	BE575706	BE575706 211961.MA	336	18	0.9	639	10	BJ108064	BJ108064
C 264	18	0.9	553	12	A2169690	SP_0114-A	337	18	0.9	641	10	BJ114124	BJ114124
C 265	18	0.9	553	10	BI236582	RE32653.5	338	18	0.9	644	10	BE273985	BE273985
C 266	18	0.9	555	10	BI228974	RE26650.5	339	18	0.9	644	12	BH442086	BH442086
C 267	18	0.9	557	12	B98152	B98152 F26D24TFB.I	340	18	0.9	645	12	A2804232	A2804232
C 268	18	0.9	558	9	AI656747	AI656747 tF53N08.x	341	18	0.9	649	9	BB160674	BB160674
C 269	18	0.9	558	10	BM108987	BM108987 EST556523	342	18	0.9	651	9	BB159570	BB159570
C 270	18	0.9	559	10	BJ111568	BJ111568 BJ111568	343	18	0.9	653	9	AM002316	AM002316
C 271	18	0.9	562	9	A0034779	A0034779 A0034779	344	18	0.9	653	10	BI215969	BI215969
C 272	18	0.9	562	10	BI375009	BI375009 RE62715.5	345	18	0.9	654	10	BI356777	BI356777
C 273	18	0.9	563	10	BE442362	BE442362 925019C05	346	18	0.9	654	12	BH091880	BH091880
C 274	18	0.9	566	9	AV610755	AV610755 AV610755	347	18	0.9	655	10	BM163651	BM163651
C 275	18	0.9	567	12	A2093968	RPTC-23-4	348	18	0.9	655	12	BM173717	BM173717
C 276	18	0.9	568	10	BE594413	BE594413 RE38396.5	349	18	0.9	659	10	BI238697	BI238697
C 277	18	0.9	572	9	AA990731	AA990731 P11_33_BO	350	18	0.9	660	10	BI374878	BI374878
C 278	18	0.9	573	10	BI237259	BI237259 RE33451.y	351	18	0.9	660	10	BM160463	BM160463
C 279	18	0.9	573	10	BE554390	ur42a05.5	352	18	0.9	661	10	BM231189	BM231189
C 280	18	0.9	575	12	BH405516	BH405516 Gm.ISB001	353	18	0.9	662	9	AV880405	AV880405
C 281	18	0.9	575	12	BH405516	BH405516 Gm.ISB001	354	18	0.9	665	12	BH615359	BH615359
C 282	18	0.9	579	10	BE506173	BE506173 AT06574.5	355	18	0.9	666	10	BI228062	BI228062
C 283	18	0.9	580	9	AA521571	AA521571 v115602.f	356	18	0.9	670	10	BI238238	BI238238
C 284	18	0.9	581	10	BI173592	BI173592 RE16542.5	357	18	0.9	670	10	BI483728	BI483728
C 285	18	0.9	581	10	BI237311	BI237311 RE33523.5	358	18	0.9	670	10	BE502379	BE502379
C 286	18	0.9	581	10	BM136167	BM136167 WHE2606.A	359	18	0.9	671	10	BM171161	BM171161
C 287	18	0.9	584	10	WB0408	WB0408 zH49f05.s1	360	18	0.9	672	10	BI485537	BI485537
C 288	18	0.9	585	10	BE523572	BE523572 33-28.Ste	361	18	0.9	675	10	BE633677	BE633677
C 289	18	0.9	585	10	BM145014	TCAP1D14	362	18	0.9	676	10	BI228209	BI228209
C 290	18	0.9	587	10	BM163467	BM163467 EST565990	363	18	0.9	677	10	BI579195	BI579195
C 291	18	0.9	588	12	A2212241	A2212241 Sheared.D	364	18	0.9	679	10	BM170305	BM170305
C 292	18	0.9	588	12	A2801390	A2801390 2M0059P13	365	18	0.9	679	12	AG059706	AG059706
C 293	18	0.9	589	9	AI719073	AI719073 as65e08.x	366	18	0.9	679	12	AG072134	AG072134
C 294	18	0.9	589	9	AI971249	AI971249 wr27a10.x	367	18	0.9	681	9	AT774508	AT774508
C 295	18	0.9	589	12	A2856098	A2856098 2M0160H19	368	18	0.9	684	10	BE635702	BE635702
C 296	18	0.9	590	9	AI760647	AI760647 w166904.x	369	18	0.9	686	9	AM969044	AM969044
C 297	18	0.9	592	10	BI232447	BI232447 RE28515.5	370	18	0.9	687	10	BJ004084	BJ004084
C 298	18	0.9	596	9	AI928200	AI928200 w095c11.x	371	18	0.9	688	10	BI231283	BI231283
C 299	18	0.9	596	10	BR491199	BR491199 AT27886.5	372	18	0.9	688	10	BMK009398	BMK009398
C 300	18	0.9	598	9	AA390638	AA390638 ID08837.5	373	18	0.9	691	9	BE208978	BE208978
C 301	18	0.9	602	10	BI376051	BI376051 RE63975.5	374	18	0.9	692	12	AG114407	AG114407
C 302	18	0.9	605	10	BG180550	BG180550 602331483	375	18	0.9	693	9	AU010111	AU010111
C 303	18	0.9	609	9	AV385599	AV385599 AV385599	376	18	0.9	695	10	BI236610	BI236610
C 304	18	0.9	609	12	A2470485	A2470485 1M0284C01	377	18	0.9	698	10	BI368144	BI368144
C 305	18	0.9	610	10	BI362389	BI362389 RE46933.5	378	18	0.9	699	10	BM407630	BM407630
C 306	18	0.9	611	10	BJ025129	BJ025129 BJ025129	379	18	0.9	700	9	AL508089	AL508089
C 307	18	0.9	611	10	BE423488	BE423488 WHE0064_G	380	18	0.9	704	10	BE491603	BE491603
C 308	18	0.9	611	12	AQ324245	AQ324245 mgxB0017F	381	18	0.9	704	10	BI231108	BI231108
C 309	18	0.9	613	12	AQ681648	AQ681648 HS_2161_A	382	18	0.9	705	12	BH058352	BH058352

383	18	0.9	706	10	B1559365	B1559365 603253036	C 456	18	0.9	995	9	AL572762	AL572762
384	18	0.9	709	10	B1484189	B1484189 RE67248.5	457	18	0.9	1014	10	BM423728	BM423728
385	18	0.9	713	10	B1522087	B1522087 603081815	458	18	0.9	1020	10	BG822861	BG822861
386	18	0.9	714	9	BE225469	BE225469 MD0377 Me	C 459	18	0.9	1036	12	CNS01G4N	CNS01G4N
387	18	0.9	714	12	BH118334	BH118334 RPCI-24-2	C 460	18	0.9	1040	12	CNS02NAX	CNS02NAX
388	18	0.9	716	10	B1215435	B1215435 RE21836.5	C 461	18	0.9	1042	9	AL517775	AL517775
389	18	0.9	716	10	BM239189	BM239189 K0530D07-	C 462	18	0.9	1049	10	BM457075	BM457075
390	18	0.9	717	9	AV401272	AV401272 AV401272	C 463	18	0.9	1050	9	AL582276	AL582276
391	18	0.9	717	9	BE071793	BE071793 RCO-BT052	C 464	18	0.9	1089	10	BG787895	BG787895
392	18	0.9	719	12	BH013921	BH013921 TPCBRT0TH	C 465	18	0.9	1100	12	CNS055AM	CNS055AM
393	18	0.9	719	12	AQ361498	AQ361498 m9XB0004K	C 466	18	0.9	1415	10	BG845569	BG845569
394	18	0.9	724	9	AW515087	AW515087 x911b07.x	C 467	18	0.9	114	12	AZ402379	AZ402379
395	18	0.9	733	12	BH570007	BH570007 BOHNB0TR	C 468	18	0.9	116	12	AQ630915	AQ630915
396	18	0.9	738	10	BF506094	BF506094 AT08477.5	C 469	18	0.9	130	10	B1996618	B1996618
397	18	0.9	743	10	BM162957	BM162957 EST565480	C 470	18	0.9	132	10	BM110457	BM110457
398	18	0.9	743	10	BF037425	BF037425 601461002	C 471	18	0.9	137	10	BF014246	BF014246
399	18	0.9	750	10	BE876638	BE876638 601488201	C 472	18	0.9	153	10	BF329139	BF329139
400	18	0.9	754	9	AL526524	AL526524 AL526524	C 473	18	0.9	163	9	AU213065	AU213065
401	18	0.9	755	10	BE875180	BE875180 601488667	C 474	18	0.9	166	10	BE757712	BE757712
402	18	0.9	763	12	BH492670	BH492670 BOHGC30TF	C 475	18	0.9	167	10	BG868566	BG868566
403	18	0.9	770	12	BH101159	BH101159 RPCI-24-3	C 476	18	0.9	173	9	AW796288	AW796288
404	18	0.9	777	9	AU000571	AU000571 AU000571	C 477	18	0.9	182	10	BE841160	BE841160
405	18	0.9	778	10	BM010510	BM010510 603631475	C 478	18	0.9	186	9	AA502132	AA502132
406	18	0.9	779	9	AL623065	AL623065 AEMTAD21	C 479	18	0.9	189	9	AA806651	AA806651
407	18	0.9	780	10	BM159621	BM159621 EST562144	C 480	18	0.9	191	9	AA422649	AA422649
408	18	0.9	783	9	AL563059	AL563059 AL563059	C 481	18	0.9	191	9	AA619734	AA619734
409	18	0.9	788	9	AV400991	AV400991 AV400991	C 482	18	0.9	197	9	AA681308	AA681308
410	18	0.9	788	9	AA567948	AA567948 H102035.5	C 483	18	0.9	197	9	AW655556	AW655556
411	18	0.9	791	10	BG760386	BG760386 602716807	C 484	18	0.9	200	9	AL850166	AL850166
412	18	0.9	793	12	BH471988	BH471988 BCGFT36TR	C 485	18	0.9	200	10	BF774407	BF774407
413	18	0.9	799	10	BG389916	BG389916 602415459	C 486	18	0.9	203	10	F15156	F15156
414	18	0.9	799	12	BH490388	BH490388 BOHBI13TR	C 487	18	0.9	206	10	BE821671	BE821671
415	18	0.9	802	10	BF968244	BF968244 602269611	C 488	18	0.9	211	10	BG017720	BG017720
416	18	0.9	815	10	BG780345	BG780345 SEAUW0000	C 489	18	0.9	213	10	T48934	T48934
417	18	0.9	823	9	AV401129	AV401129 AV401129	C 490	18	0.9	215	9	BM145109	BM145109
418	18	0.9	835	9	AL572467	AL572467 AL572467	C 491	18	0.9	217	9	AI960801	AI960801
419	18	0.9	841	9	AL579300	AL579300 AL579300	C 492	18	0.9	218	10	BF546288	BF546288
420	18	0.9	841	12	AG137000	AG137000 Pan Trogl	C 493	18	0.9	220	9	AV335016	AV335016
421	18	0.9	844	12	BH023317	BH023317 GH-MBB000	C 494	18	0.9	224	9	AA566015	AA566015
422	18	0.9	844	12	BH511647	BH511647 BCGFT71TF	C 495	18	0.9	226	10	B1316911	B1316911
423	18	0.9	846	10	BE972796	BE972796 601652028	C 496	18	0.9	227	10	B1435280	B1435280
424	18	0.9	849	10	BF527018	BF527018 602070612	C 497	18	0.9	234	10	BE484532	BE484532
425	18	0.9	849	12	BH583147	BH583147 BCGMT35TF	C 498	18	0.9	235	10	H21312	H21312
426	18	0.9	854	12	AQ450136	AQ450136 500008D02	C 499	18	0.9	236	10	BE725065	BE725065
427	18	0.9	856	10	BE797809	BE797809 601586425	C 500	18	0.9	240	9	AU074507	AU074507
428	18	0.9	858	10	BG623369	BG623369 602648413	C 501	18	0.9	243	12	AZ258446	AZ258446
429	18	0.9	867	9	AL574257	AL574257 AL574257	C 502	18	0.9	244	9	AV273323	AV273323
430	18	0.9	869	10	BG570663	BG570663 602590965	C 503	18	0.9	244	10	BE527808	BE527808
431	18	0.9	871	10	BM006012	BM006012 603613748	C 504	18	0.9	246	9	AA570408	AA570408
432	18	0.9	872	10	B1522040	B1522040 603081736	C 505	18	0.9	247	9	AW706163	AW706163
433	18	0.9	874	9	AL571869	AL571869 AL571869	C 506	18	0.9	249	9	AI556653	AI556653
434	18	0.9	877	10	BG476325	BG476325 602521818	C 507	18	0.9	256	9	BB579437	BB579437
435	18	0.9	881	12	BH160577	BH160577 EMT0011TR	C 508	18	0.9	256	9	BB597624	BB597624
436	18	0.9	884	10	BG330475	BG330475 602430147	C 509	18	0.9	257	9	AV163047	AV163047
437	18	0.9	902	9	AL550497	AL550497 AL550497	C 510	18	0.9	261	9	BB010222	BB010222
438	18	0.9	906	9	AL580498	AL580498 AL580498	C 511	18	0.9	262	10	B1015339	B1015339
439	18	0.9	908	12	BM010512	BM010512 603631477	C 512	18	0.9	264	9	BB748332	BB748332
440	18	0.9	908	12	CNS05B19	CNS05B19 Teltrodon	C 513	18	0.9	266	10	BF220401	BF220401
441	18	0.9	909	10	BG423508	BG423508 602449370	C 514	18	0.9	268	9	BA466539	BA466539
442	18	0.9	910	12	CNS036T1	AL230446 Teltrodon	C 515	18	0.9	268	10	B1922332	B1922332
443	18	0.9	911	12	CNS06J3K	AL403686 T end of	C 516	18	0.9	272	9	AV300066	AV300066
444	18	0.9	912	10	B1078938	B1078938 602873184	C 517	18	0.9	272	9	BB195297	BB195297
445	18	0.9	912	10	BE743286	BE743286 601574885	C 518	18	0.9	274	9	BB195297	BB195297
446	18	0.9	929	10	BF791469	BF791469 602251511	C 519	18	0.9	277	9	BB351790	BB351790
447	18	0.9	945	10	B1089149	B1089149 602854004	C 520	18	0.9	277	9	BB722075	BB722075
448	18	0.9	945	10	BE617956	BE617956 601442184	C 521	18	0.9	286	9	AW568359	AW568359
449	18	0.9	947	9	AL579071	AL579071 AL579071	C 522	18	0.9	286	10	BF186611	BF186611
450	18	0.9	949	9	BE213916	BE213916 HV-CEB000	C 523	18	0.9	289	10	BG239086	BG239086
451	18	0.9	962	12	CNS07DB3	AL440245 T end of	C 524	18	0.9	289	12	B06423	B06423
452	18	0.9	970	12	CNS04AMW	AL282425 Teltrodon	C 525	18	0.9	290	9	BE180963	BE180963
453	18	0.9	975	9	AL572942	AL572942 AL572942	C 526	18	0.9	290	10	BE477329	BE477329
454	18	0.9	978	10	BM424224	BM424224 AGENCOURT	C 527	18	0.9	291	9	BB194398	BB194398
455	18	0.9	978	10	BE740965	BE740965 601595328	C 528	18	0.9	294	9	AV293940	AV293940

529	17	0.9	295	10	BE824216	BE824216 GM700023A	602	17	0.9	370	9	AU204179	AU204179
530	17	0.9	296	10	BF391767	BF391767 UI-R-CA1-	603	17	0.9	371	12	A2127658	A2127658 OSUNB009
531	17	0.9	299	9	AM334777	AM334777 S378 ACS	604	17	0.9	370	10	BE802196	BE802196 ST25F05.Y
532	17	0.9	300	9	AA229159	AA229159 nc45B11.s	605	17	0.9	371	10	AV198561	AV198561 AV198561
533	17	0.9	300	9	BA400928	BA400928	606	17	0.9	375	10	CA3941	CA3941 CA3941 Yuj1
534	17	0.9	302	9	AV167272	AV167272	607	17	0.9	376	12	AQ045322	AQ045322 RPTC111-36
535	17	0.9	302	9	BB381934	BB381934	608	17	0.9	377	9	AW788524	AW788524 C00516-F
536	17	0.9	303	10	BE317699	BE317699 NF053F03L	609	17	0.9	377	9	BB804538	BB804538
537	17	0.9	304	10	BF964473	BF964473 RCI-NN023	610	17	0.9	377	10	H89049	H89049 yw24c02.r1
538	17	0.9	304	10	BF054642	BF054642 EST439872	611	17	0.9	378	10	BG159315	BG159315 OV2.6.C01
539	17	0.9	307	9	BB838179	BB838179	612	17	0.9	378	10	B1359863	B1359863 384606.MA
540	17	0.9	308	10	BF958519	BF958519 OV2-NN004	613	17	0.9	378	10	CA48533	CA48533 C48533 Yuj1
541	17	0.9	310	9	AA793720	AA793720 vr95g10.r	614	17	0.9	378	12	A2493775	A2493775 1M0328E08
542	17	0.9	311	9	AA947333	AA947333 Oq49508.s	615	17	0.9	379	12	AC010574	AC010574
543	17	0.9	311	10	B1033010	B1032010 PMO-MT042	616	17	0.9	380	10	CA7511	CA7511 CA7511 Yuj1
544	17	0.9	311	10	BT032021	BT032021 PMO-MT042	617	17	0.9	381	12	AO061280	AO061280 CIT-HSP-2
545	17	0.9	312	9	AW347101	AW347101 30235.MAR	618	17	0.9	382	9	A1923125	A1923125 wM83n06.x
546	17	0.9	315	10	UB5028	UB5028 DM085028.Dr	619	17	0.9	382	12	AQ959102	AQ959102 LERREH30TR
547	17	0.9	319	10	B1316939	B1316939 satf82e07.	620	17	0.9	383	9	AA928151	AA928151 ON86F01.s
548	17	0.9	320	9	AM659554	AM659554 96860.MAR	621	17	0.9	383	12	AO967181	AO967181 LERIO14TF
549	17	0.9	320	10	H11416	H11416 yml13q05.r1	622	17	0.9	384	12	AM543420	AM543420 C0169F10-
550	17	0.9	321	10	BE940989	BE940989 EST420568	623	17	0.9	384	12	AO804601	AO804601 HS-2263_B
551	17	0.9	322	9	AM484678	AM484678 62094.MAR	624	17	0.9	385	10	T47128	T47128 yb52901.r1
552	17	0.9	323	9	BB100939	BB100939	625	17	0.9	386	9	BB804917	BB804917
553	17	0.9	325	10	BG158832	BG158832 RH122_43_	626	17	0.9	386	12	BH126745	BH126745 BARC-Salt
554	17	0.9	330	10	AA366781	AA366781 EST77771	627	17	0.9	388	10	BG644749	BG644749 EST506368
555	17	0.9	330	10	B1372506	B1372506 RE59481.5	628	17	0.9	388	12	BH016030	BH016030 TDGCH56TH
556	17	0.9	330	10	BT398532	BT398532 952002E01	629	17	0.9	389	9	AM987158	AM987158 w15n05.x
557	17	0.9	334	12	CNS01RZA	AL164575 Telradon	630	17	0.9	389	12	A2688866	A2688866 2M0180E13
558	17	0.9	335	12	A2849782	A2849782 2M0151B24	631	17	0.9	390	10	B1926856	B1926856 EST546745
559	17	0.9	336	12	A2924995	A2924995 4910.ez31	632	17	0.9	390	12	A2752618	A2752618 RPTC-24-1
560	17	0.9	337	9	AM217308	AM217308 EST296027	633	17	0.9	391	9	AM450960	AM450960 UI-H-B13-
561	17	0.9	338	9	AM569761	AM569761 s181b01.Y	634	17	0.9	392	9	AL635835	AL635835
562	17	0.9	339	10	C30724	C30724 HMG800475	635	17	0.9	392	12	BH098871	BH098871 RPTC-24-3
563	17	0.9	340	9	A1363761	A1363761 qv79802.x	636	17	0.9	393	12	B30710	B30710 HS-1003-A1-
564	17	0.9	341	10	F20510	F20510 HSPD04860.H	637	17	0.9	394	10	BG134036	BG134036 EST466928
565	17	0.9	342	9	AU176520	AU176520	638	17	0.9	397	9	AA558145	AA558145 n126d10.s
566	17	0.9	344	10	BM442866	BM442866 EBR002.SQ	639	17	0.9	398	9	A1273733	A1273733 q147606.x
567	17	0.9	345	10	B134799	B134799	640	17	0.9	398	10	BG927311	BG927311 HNC64-1-C
568	17	0.9	345	12	AQ756347	AQ756347 HS-5210_B	641	17	0.9	399	9	AA386929	AA386929
569	17	0.9	346	9	AV197804	AV197804	642	17	0.9	399	10	N23811	N23811 yx36e10.r1
570	17	0.9	346	9	AA400475	AA400475 zu64e10.s	643	17	0.9	399	12	AQ121279	AQ121279 HS-3075_A
571	17	0.9	346	10	T05409	T05409 EST03298.Fe	644	17	0.9	400	10	BE571388	BE571388 601330227
572	17	0.9	346	10	BF609597	BF609597 NXSI_047_	645	17	0.9	401	9	AM086060	AM086060 xc77g03.x
573	17	0.9	347	12	NS0424	NS0424 YYR88408.s1	646	17	0.9	402	10	BF526719	BF526719 602070559
574	17	0.9	347	12	A2776537	A2776537 2M0010C18	647	17	0.9	403	9	AA105576	AA105576
575	17	0.9	348	10	B1934423	B1934423 EST554312	648	17	0.9	403	12	BH140250	BH140250
576	17	0.9	350	9	BH004658	BH004658 BMAC08G0	649	17	0.9	404	10	w97823	w97823
577	17	0.9	350	9	AU112203	AU112203	650	17	0.9	405	9	AM008231	AM008231 wv52a02.x
578	17	0.9	351	9	A1825450	A1825450 wb73b01.x	651	17	0.9	405	10	BG137180	BG137180 EST477622
579	17	0.9	351	9	AU112508	AU112508	652	17	0.9	406	9	BE220007	BE220007 hv65n05.x
580	17	0.9	351	10	BE756537	BE756537 210705.MA	653	17	0.9	407	9	A1821281	A1821281 ac73n01.x
581	17	0.9	351	10	BE850184	BE850184 uw13h11.Y	654	17	0.9	407	9	A1945739	A1945739
582	17	0.9	352	10	BF044171	BF044171	655	17	0.9	407	10	BG012541	BG012541 ILS-GN023
583	17	0.9	355	9	AV011169	AV011169	656	17	0.9	408	10	BF774413	BF774413 284079.MA
584	17	0.9	355	10	B1972374	B1972374 sag90f01.	657	17	0.9	408	12	B72961	B72961 RPTC11-10M1
585	17	0.9	357	9	AM085070	AM085070 xe05c01.x	658	17	0.9	409	10	BF939304	BF939304 nad6b04.
586	17	0.9	358	9	AM614412	AM614412 hg90g10.x	659	17	0.9	409	10	B1112593	B1112593
587	17	0.9	360	9	AU098178	AU098178	660	17	0.9	410	10	BT039948	BT039948 MR4-NT014
588	17	0.9	360	9	AV194655	AV194655	661	17	0.9	411	10	BF188721	BF188721 EST445008
589	17	0.9	360	10	C50400	C50400	662	17	0.9	412	9	AM662635	AM662635 h133n04.x
590	17	0.9	360	10	D69839	D69839 CELK092G4F	663	17	0.9	414	9	AM519964	AM519964 660055H10
591	17	0.9	360	10	D73983	D73983 CELK067D2F	664	17	0.9	416	9	AA461730	AA461730 BP230006A
592	17	0.9	360	10	D74063	D74063 CELK074E1F	665	17	0.9	418	10	BM286810	BM286810 527013.MA
593	17	0.9	360	10	D74690	D74690 CELK084B5F	666	17	0.9	418	12	AQ026779	AQ026779 HS-3241_A
594	17	0.9	360	12	AQ283182	AQ283182 RPTC111-80	667	17	0.9	419	10	BF052856	BF052856 EST438086
595	17	0.9	365	9	AM928766	AM928766 EST337554	668	17	0.9	419	10	BF705726	BF705726 248770.MA
596	17	0.9	368	10	BM217713	BM217713 C0901B02-	669	17	0.9	420	9	A1440690	A1440690 sa67g12.Y
597	17	0.9	368	10	BM272236	BM272236 1q42a09.x	670	17	0.9	421	10	BF071522	BF071522 st62b10.Y
598	17	0.9	368	10	R04076	R04076 PK19a12.s1	671	17	0.9	422	9	AA830935	AA830935 oc58a09.s
599	17	0.9	369	10	BG273315	BG273315 OV2.25.H0	672	17	0.9	422	9	AW754948	AW754948 PC08G02.P
600	17	0.9	369	10	BE520504	BE520504 M13A10XP	673	17	0.9	422	10	BM390569	BM390569 UI-R-CM1-
601	17	0.9	369	10	BF501010	BF501010 AT16194.5	674	17	0.9	422	10	BE673931	BE673931 7d71e08.x

c 675	17	0.9	422	10	BE862173	BE862173	UT-M-BH0-	c 748	17	0.9	460	12	AQ0920491	AQ0920491	RPCT-23-2
c 676	17	0.9	423	9	BB789456	BB789456	PFESToa2	c 749	17	0.9	461	10	A1237508	A1237508	ESP234070
c 677	17	0.9	425	10	BI814528	BI814528	PFESToa2	c 750	17	0.9	461	9	C95943	C95943	Marc
c 678	17	0.9	425	10	BM312803	BM312803	1g79c10.x	c 751	17	0.9	462	12	B64101	B64101	CIT-HSP-176
c 679	17	0.9	425	12	AQ268118	AQ268118	RPCT11-73	c 752	17	0.9	463	10	BE673836	BE673836	7d70004.x
c 680	17	0.9	426	12	A2154883	A2154883	SP_0032_A	c 753	17	0.9	463	10	BE923951	BE923951	EST427720
c 681	17	0.9	427	12	AQ967751	AQ967751	LERI069TF	c 754	17	0.9	463	12	AQ332688	AQ332688	HS_5004_A
c 682	17	0.9	428	9	BB205943	BB205943	BR205943	c 755	17	0.9	464	10	AA016250	AA016250	ze33c01.s
c 683	17	0.9	429	9	BB776475	BB776475	BR776475	c 756	17	0.9	464	10	BF389741	BF389741	UT-R-BS2-
c 684	17	0.9	430	9	AM687133	AM687133	NFO06D02R	c 757	17	0.9	464	12	AQ296005	AQ296005	HS_3090_A
c 685	17	0.9	431	9	AM896747	AM896747	CNO-NN005	c 758	17	0.9	465	9	AJ396822	AJ396822	A1396822
c 686	17	0.9	431	10	BF938179	BF938179	fm73d10..y	c 759	17	0.9	466	12	A2509383	A2509383	1M0352L17
c 687	17	0.9	432	9	AA976607	AA976607	og03d02.s	c 760	17	0.9	467	9	BB759722	BB759722	BR759722
c 688	17	0.9	432	9	A1854874	A1854874	UT-M-BH0-	c 761	17	0.9	467	12	AQ635993	HS_2153_B	AQ635993
c 689	17	0.9	432	10	BG650317	BG650317	sad05B04.	c 762	17	0.9	468	12	AQ385933	RPCT11-14	AQ385933
c 690	17	0.9	432	12	AQ183615	AQ183615	HS_3197.B	c 763	17	0.9	469	9	AA024988	AA024988	9s
c 691	17	0.9	433	10	BF964510	BF964510	RC1-NN023	c 764	17	0.9	469	12	AQ936787	UNB-311R	AQ936787
c 692	17	0.9	433	12	AZ002333	AZ002333	RPCT-23-3	c 765	17	0.9	470	9	BB465385	BB465385	BB465385
c 693	17	0.9	434	9	AM489072	AM489072	UI-M-BH3-	c 766	17	0.9	470	12	A2724080	RPCT-24-7	A2724080
c 694	17	0.9	434	10	BE501780	BE501780	hw35D02.x	c 767	17	0.9	471	9	AL655394	AL655394	AL655394
c 695	17	0.9	436	10	BM434078	BM434078	1RT02B05	c 768	17	0.9	471	9	AM515751	hdb7h10.x	hdb7h10.x
c 696	17	0.9	436	12	BM170059	BM170059	SALK_0023	c 769	17	0.9	471	9	AM515751	hdb7h10.x	hdb7h10.x
c 697	17	0.9	437	10	BI321307	BI321307	SAF50C10.	c 770	17	0.9	471	9	BB657258	BB657258	BB657258
c 698	17	0.9	439	10	BF504887	BF504887	AT06979..5	c 771	17	0.9	471	10	BF746041	IL2-BT080	BF746041
c 699	17	0.9	440	12	AQ447347	AQ447347	mgxb0006L	c 772	17	0.9	472	9	AM465238	BP230018A	AM465238
c 700	17	0.9	441	9	A1285705	A1285705	qu55h08.x	c 773	17	0.9	472	9	AL376980	AL376980	AL376980
c 701	17	0.9	441	12	AQ451638	AQ451638	HS_5157_A	c 774	17	0.9	473	12	AQ787765	HS_2025_B	AQ787765
c 702	17	0.9	442	9	AM791729	AM791729	D00771-R	c 775	17	0.9	473	12	AQ672419	HS_2276_B	AQ672419
c 703	17	0.9	442	9	BB710569	BB710569	BR710569	c 776	17	0.9	474	10	BI847276	469071.MA	BI847276
c 704	17	0.9	442	12	BH273263	BH273263	CH230-165	c 777	17	0.9	474	10	BI847276	469071.MA	BI847276
c 705	17	0.9	443	9	A1429807	A1429807	mg06b07.x	c 778	17	0.9	475	9	AA287260	z49904.r	AA287260
c 706	17	0.9	443	9	A1582628	A1582628	tn47902.x	c 779	17	0.9	477	9	A1458669	A1458669	TK13c11.x
c 707	17	0.9	443	9	AM183674	AM183674	x186e04.x	c 780	17	0.9	477	10	BE646458	7e87B06.x	BE646458
c 708	17	0.9	443	10	BF925658	BF925658	CME-NT019	c 781	17	0.9	478	10	BM454904	AGENCOURT	BM454904
c 709	17	0.9	445	9	AA419503	AA419503	zvo3b03.r	c 782	17	0.9	479	10	BC011415	R1-057	BC011415
c 710	17	0.9	445	10	BM252558	BM252558	SmEST4a40	c 783	17	0.9	479	10	BE665653	150057.MA	BE665653
c 711	17	0.9	447	9	A1459733	A1459733	ap16110.x	c 784	17	0.9	479	10	BF560112	UT-R-A1-e	BF560112
c 712	17	0.9	447	10	BF412152	BF412152	UI-R-BT1-	c 785	17	0.9	480	10	R17027	yf45c02.r2	R17027
c 713	17	0.9	447	12	AZ089401	AZ089401	RPCT-23-4	c 786	17	0.9	480	10	RS9405	h17a02.r1	RS9405
c 714	17	0.9	449	9	AM500820	AM500820	UI-HF-BP0	c 787	17	0.9	481	12	AQ568041	HS_2135_A	AQ568041
c 715	17	0.9	449	10	BG000024	BG000024	RC3-GN004	c 788	17	0.9	481	9	AA769665	ob19b06.s	AA769665
c 716	17	0.9	450	9	AA701527	AA701527	z167b01.s	c 789	17	0.9	482	9	AA725832	a135c08.s	AA725832
c 717	17	0.9	450	9	AL376981	AL376981	MEB27E05	c 790	17	0.9	482	9	AM310039	s130a02.x	AM310039
c 718	17	0.9	450	9	AU215814	AU215814	AU215814	c 791	17	0.9	482	10	R00005	ye70e05.s1	R00005
c 719	17	0.9	450	9	BB726139	BB726139	BB726139	c 792	17	0.9	483	12	BH363747	CH230-190	BH363747
c 720	17	0.9	450	10	BJ201317	BJ201317	BJ201317	c 793	17	0.9	483	9	A1911582	tv73c09.x	A1911582
c 721	17	0.9	450	10	C75413	C75413	Humma	c 794	17	0.9	483	12	A2171404	SP_0119_B	A2171404
c 722	17	0.9	450	10	BF481446	BF481446	FWL_19_CO	c 795	17	0.9	484	9	AM734380	sk19d07.y	AM734380
c 723	17	0.9	451	9	AA883061	AA883061	am17h01.s	c 796	17	0.9	485	10	BG047424	sa871d07.y	BG047424
c 724	17	0.9	451	10	BI511261	BI511261	BI160005A	c 797	17	0.9	485	10	BM303173	SmEST4a45	BM303173
c 725	17	0.9	451	10	BI683351	BI683351	464703.MA	c 798	17	0.9	486	9	AM399423	EST309923	AM399423
c 726	17	0.9	451	10	BF413404	BF413404	UT-R-CA0-	c 799	17	0.9	486	12	AQ293324	HS_2239_B	AQ293324
c 727	17	0.9	452	9	AL385832	AL385832	MEBC31A04	c 800	17	0.9	487	9	A1431254	at56d01.x	A1431254
c 728	17	0.9	452	9	AM583728	AM583728	1a03906.y	c 801	17	0.9	488	9	AA126244	z186f01.s	AA126244
c 729	17	0.9	453	9	AA797470	AA797470	vw27908.r	c 802	17	0.9	488	9	A1417809	tg71f06.x	A1417809
c 730	17	0.9	454	10	BG726978	BG726978	sa829f01.	c 803	17	0.9	488	10	BF774053	263599.MA	BF774053
c 731	17	0.9	454	12	AQ114911	AQ114911	CIT-HSP-2	c 804	17	0.9	488	10	BF490983	AR27630.5	BF490983
c 732	17	0.9	454	12	AQ579927	AQ579927	t135463b	c 805	17	0.9	489	10	BE955178	P1146_H1	BE955178
c 733	17	0.9	455	9	A1946323	A1946323	bs23f02.y	c 806	17	0.9	490	10	BI348474	d8g34b05.	BI348474
c 734	17	0.9	455	10	BE461632	BE461632	EST413051	c 807	17	0.9	490	10	BE448544	ut92c06.y	BE448544
c 735	17	0.9	456	9	A1468722	A1468722	th85e07.x	c 808	17	0.9	491	9	AL044668	DKF2p434B	AL044668
c 736	17	0.9	456	10	W61379	W61379	zdt7g10..r1	c 809	17	0.9	491	10	BG557557	EWL_53_AO	BG557557
c 737	17	0.9	456	10	BE754369	BE754369	207864.MA	c 810	17	0.9	491	10	BE863644	UT-M-BH0-	BE863644
c 738	17	0.9	456	12	AQ912330	AQ912330	nbe00035f	c 811	17	0.9	491	12	A2026875	RPCT-23-3	A2026875
c 739	17	0.9	457	9	AM426132	AM426132	59675.MAR	c 812	17	0.9	492	9	A1854627	ut43d04.x	A1854627
c 740	17	0.9	457	9	AM469541	AM469541	hd19e02.x	c 813	17	0.9	492	9	AW909105	u1f3d04.x	AW909105
c 741	17	0.9	458	9	A1138392	A1138392	q8d4a04.x	c 814	17	0.9	492	10	BF986233	RC3-GN004	BF986233
c 742	17	0.9	458	9	A1218701	A1218701	ok39b05.x	c 815	17	0.9	493	9	AA177702	mt33f11.r	AA177702
c 743	17	0.9	459	12	AZ180126	AZ180126	SP_0167_B	c 816	17	0.9	493	10	BF557292	UT-R-C0-9	BF557292
c 744	17	0.9	460	9	BB670325	BB670325	BB670325	c 817	17	0.9	493	12	AQ005928	CIT-HSP-2	AQ005928
c 745	17	0.9	460	9	BB746182	BB746182	BR746182	c 818	17	0.9	494	12	AQ139856	HS_3104_A	AQ139856
c 746	17	0.9	460	10	BI936150	BI936150	PFESToa2	c 819	17	0.9	494	10	BF389658	UT-R-BS2-	BF389658
c 747	17	0.9	460	12	AQ676740	AQ676740	HS_2146_B	c 820	17	0.9	494	12	AQ380944	RPCT11-16	AQ380944

821	17	0.9	495	10	B1699905	B1699905 saq49e03.	894	17	0.9	526	12	A0252267	A0252267 HS_3194_A
c 822	17	0.9	496	9	AL385833	AL385833 MTRC31A04	895	17	0.9	527	10	AM886339	AM886339 MR2-OT007
823	17	0.9	496	9	AM462495	BP230009B	c 896	17	0.9	527	10	C84415	C84415 C84415 oste
824	17	0.9	496	10	BP592986	BP592986 7j97h04.x	c 897	17	0.9	527	12	A2282449	A2282449 RPTC11-15
825	17	0.9	496	12	A0550483	AO550483 RPTC11-14	c 898	17	0.9	527	12	A0374965	A0374965 RPTC11-15
826	17	0.9	499	9	AA176323	AA176323 zp08g04.x	c 899	17	0.9	528	10	BF173998	BF173998 OSTF010B2
c 827	17	0.9	500	9	AM680656	WS1_6_H03	c 900	17	0.9	528	10	BF259663	BF259663 HVSMEF001
828	17	0.9	501	10	BG553760	BG553760 dac2e07.	c 901	17	0.9	529	9	AM408393	AM408393 UI-HF-BK0
c 829	17	0.9	501	10	B146155	B146155 BJT46155	c 902	17	0.9	529	10	BF997009	BF997009 OV3-GN020
c 830	17	0.9	501	12	CNS02WSB	AT217460 Tetracodon	c 903	17	0.9	529	10	R30162	R30162 12767 Lambd
831	17	0.9	502	10	BR297233	BR297233 049PBA11	c 904	17	0.9	529	10	BE941026	BE941026 EST420605
832	17	0.9	503	9	AW018563	AW018563 fd48e04.y	c 905	17	0.9	529	12	AQ767775	AQ767775 HS_3086_B
c 833	17	0.9	503	12	A2801392	A2801392 2M0059P15	c 906	17	0.9	530	9	AL644823	AL644823 AL644823
c 834	17	0.9	504	9	AV712647	AV712647 AV712647	c 907	17	0.9	530	10	AU207564	AU207564 AU207564
c 835	17	0.9	504	12	A2925905	A2925905 476.d1e04	c 908	17	0.9	530	12	A2247278	A2247278 RPTC1-23-1
c 836	17	0.9	504	12	B89564	B89564 RPTC11-2013	c 909	17	0.9	531	12	BH377386	BH377386 AG-ND-121
c 837	17	0.9	505	10	AM329281	AM329281 N200501e	c 910	17	0.9	532	9	AT056100	AT056100 OW33812.S
c 838	17	0.9	505	10	BM094728	BM094728 sa320c04.	c 911	17	0.9	532	9	AM398542	AM398542 EST309042
839	17	0.9	505	12	AQ397460	AQ397460 mqxb0017L	c 912	17	0.9	532	9	AM566982	AM566982 FK30a12.x
840	17	0.9	506	12	A2160110	SP_0065_A	c 913	17	0.9	533	10	BF193826	BF193826 245534 MA
841	17	0.9	506	12	AQ497913	HS_5080_B	c 914	17	0.9	533	12	AQ980183	AQ980183 RPTC1-23-3
c 842	17	0.9	507	10	B1704925	B1704925 ft59f10.y	c 915	17	0.9	533	12	A2802654	A2802654 2M0061F07
843	17	0.9	508	10	BJ098531	BJ098531 BJ098531	c 916	17	0.9	534	9	A1763361	A1763361 w155f08.x
844	17	0.9	509	9	AM267877	AM267877 EST306155	c 917	17	0.9	534	10	BJ102562	BJ102562 BJ102562
845	17	0.9	509	10	BF947219	BF947219 MR3-NM021	c 918	17	0.9	534	10	BM187885	BM187885 sa3182b09.
c 846	17	0.9	509	10	BM309223	BM309223 sak56b04.	c 919	17	0.9	534	12	A0808413	A0808413 HS_4553_B
847	17	0.9	509	10	R61011	R61011 yM05a04.r1	c 920	17	0.9	535	9	AU206291	AU206291 AU206291
848	17	0.9	509	10	BF039684	BF039684 BP250007B	c 921	17	0.9	535	9	AV667244	AV667244 AV667244
c 849	17	0.9	509	12	A0885646	HS_5525_A	c 922	17	0.9	535	9	AM284663	AM284663 LG1_271_H
c 850	17	0.9	510	9	AU205815	AU205815 AU205815	c 923	17	0.9	535	9	AM706275	AM706275 s154d01.y
851	17	0.9	510	10	BJ040518	BJ040518 BJ040518	c 924	17	0.9	535	9	BM756760	BM756760 BB756760
c 852	17	0.9	511	10	AM058563	AM058563 wx3e04.x	c 925	17	0.9	536	9	AM658484	AM658484 94545 MAR
c 853	17	0.9	511	10	N25568	YX76h08.s1	c 926	17	0.9	538	10	AT167363	AT167363 w195f06.y
c 854	17	0.9	512	12	A2926957	A2926957 476.d1s13	c 927	17	0.9	538	10	BC694487	BC694487 355133 MA
c 855	17	0.9	512	12	AQ050523	AQ050523 RPTC1-11-2	c 928	17	0.9	538	10	BF193753	BF193753 245443 MA
c 856	17	0.9	512	12	AQ0584781	RPCT11-14	c 929	17	0.9	538	12	AQ0959103	AQ0959103 LERH30TR
857	17	0.9	513	9	AA272964	AA272964 SW31CA245	c 930	17	0.9	538	12	AQ0281956	AQ0281956 RPTC11-180
c 858	17	0.9	513	9	AM988748	u120f01.y	c 931	17	0.9	539	10	BF422810	BF422810 HC.d11.17
c 859	17	0.9	513	10	BP076187	BP076187 225692.MA	c 932	17	0.9	539	12	A2552724	A2552724 RPTC1-23-2
860	17	0.9	513	10	BF615176	BF615176 dg20e08.y	c 933	17	0.9	540	10	BM127656	BM127656 1e98b03.y
861	17	0.9	513	12	A0565251	HS_5355_B	c 934	17	0.9	540	10	BF390812	BF390812 UI-R-CA1
c 862	17	0.9	516	12	N25618	N25618 YX77h06.s1	c 935	17	0.9	541	10	BG816351	BG816351 dac10g06.
c 863	17	0.9	516	12	AQ407500	HS_5092_B	c 936	17	0.9	541	10	BF297067	BF297067 046bpf12
864	17	0.9	517	9	AV602989	AV602989 AV602989	c 937	17	0.9	541	12	A2164893	A2164893 SP_0078_A
c 865	17	0.9	517	9	AM973416	EST85561	c 938	17	0.9	542	9	AM493535	AM493535 UI-M-BH3
866	17	0.9	517	10	B1258495	602972d91	c 939	17	0.9	542	10	BE601003	BE601003 P11_95-GO
c 867	17	0.9	517	10	BE426082	WHE0328_H	c 940	17	0.9	542	9	AM473439	AM473439 hb04g03.x
868	17	0.9	517	10	BE922834	BE922834 EST426603	c 941	17	0.9	543	10	BE600388	BE600388 P11_95-GO
c 869	17	0.9	518	10	BE918644	FMI_2_G09	c 942	17	0.9	543	10	BT701995	BT701995 sag40912.
c 870	17	0.9	519	10	BJ100918	BJ100918 BJ100918	c 943	17	0.9	544	10	BM218613	BM218613 C0913D04
871	17	0.9	519	10	BE834013	MR2-OTM07	c 944	17	0.9	544	10	BE753081	BE753081 205971.MA
c 872	17	0.9	519	12	AQ416933	RPTC1-11-1	c 945	17	0.9	545	10	BF009491	BF009491 ss91e01.y
873	17	0.9	520	9	AA673633	vo57e02.x	c 946	17	0.9	545	10	AQ985651	AQ985651 RPTC1-20-3
c 874	17	0.9	520	9	AA680311	ac86f04.s	c 947	17	0.9	546	10	BM364152	BM364152 BS300090
875	17	0.9	520	9	AU198744	AU198744	c 948	17	0.9	546	12	BH057210	BH057210 RPTC1-24-3
c 876	17	0.9	520	9	AU199013	AU199013 AU199013	c 949	17	0.9	546	12	BH770679	BH770679 BB770679
877	17	0.9	520	10	BG917281	602816560	c 950	17	0.9	548	9	AA588144	AA588144 nm99b06.s
c 878	17	0.9	520	12	A2927904	476.d1s12	c 951	17	0.9	548	12	AO763452	AO763452 HS_3232_B
879	17	0.9	521	9	AI264648	qK04g06.x	c 952	17	0.9	548	10	AT971613	AT971613 wT06a12.x
c 880	17	0.9	521	10	BG381854	296739.MA	c 953	17	0.9	549	9	AV384359	AV384359 AV384359
c 881	17	0.9	521	10	BJ110928	BJ110928	c 954	17	0.9	549	12	AQ375927	AQ375927 RPTC11-14
c 882	17	0.9	522	9	AM293714	UI-H-B12-	c 955	17	0.9	550	9	AM655940	AM655940 107919.MA
c 883	17	0.9	522	10	B1786935	sa15509.	c 956	17	0.9	551	10	BE628003	BE628003 uu26f10.x
c 884	17	0.9	523	9	BE219271	hv50f11.x	c 957	17	0.9	551	10	BE663863	BE663863 147899.MA
c 885	17	0.9	523	10	BU065774	BU065774	c 958	17	0.9	552	10	BI451126	BI451126 kx11a07.y
886	17	0.9	523	12	B95982	F21D8TFC IG	c 959	17	0.9	553	10	BM106579	BM106579 510486.MA
c 887	17	0.9	524	9	AI863428	AI863428 tz48C01.x	c 960	17	0.9	553	10	BM110804	BM110804 EST558340
c 888	17	0.9	524	10	BF481583	FMI_21_H0	c 961	17	0.9	553	10	BE751439	BE751439 203387.MA
c 889	17	0.9	525	10	BT063036	tl3-UTM01	c 962	17	0.9	553	10	BF193277	BF193277 244802.MA
890	17	0.9	525	10	BF076051	225447.MA	c 963	17	0.9	553	12	A0735239	A0735239 HS_3225_B
c 891	17	0.9	525	10	A2632378	1M0487A06	c 964	17	0.9	553	10	AQ252267	AQ252267 HS_3194_A
c 892	17	0.9	525	12	A2632378	1M0487A06	c 965	17	0.9	553	10	AM886339	AM886339 MR2-OT007
893	17	0.9	526	10	BE809111	BE809111 214481.MA	c 966	17	0.9	553	12	A0735239	A0735239 HS_3225_B

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	602460969F1 NIH_MGC_20	Homo sapiens cDNA clone IMAGE:457753 5'	602460969F1	1	GI:13141210	human.	Homo sapiens	1 (bases 1 to 928)	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	NIH-MGC http://mgc.nci.nih.gov/	Unpublished (1999)	Contact: Robert Strausberg, Ph. D. Email: cgabs-remail.nih.gov Tissue Procurement: ATCC/DC/DTP CDNA Library Preparation: Ling Hong/Rubin Laboratory DNA Sequencing by: Inceye Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov plate: L10M1292 row: 1 column: 18 High quality sequence stop: 747.

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
2	602460969F1 NIH_MGC_20	Homo sapiens cDNA clone IMAGE:457753 5'	602460969F1	1	GI:13141210	human.	Homo sapiens	1 (bases 1 to 928)	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	NIH-MGC http://mgc.nci.nih.gov/	Unpublished (1999)	Contact: Robert Strausberg, Ph. D. Email: cgabs-remail.nih.gov Tissue Procurement: ATCC/DC/DTP CDNA Library Preparation: Ling Hong/Rubin Laboratory DNA Sequencing by: Inceye Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov plate: L10M1292 row: 1 column: 18 High quality sequence stop: 747.

FEATURES

source

1. 928

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:457753"

/clone_lib="NIH_MGC_20"

/tissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pOTB7; Site: 1; XhoI: Site 2; EcoRI: cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

234 a 245 c 302 g 147 t

BASE COUNT

Query Match

Best Local Similarity 100.0%; Pred. No. 1.1;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

570 CTGCTGAACAGTATGAGAAAGCA 592

661 CTGCTGAACAGTATGAGAAAGCA 683

RESULT

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
2	602460969F1 NIH_MGC_20	Homo sapiens cDNA clone IMAGE:457753 5'	602460969F1	1	GI:13141210	human.	Homo sapiens	1 (bases 1 to 928)	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	NIH-MGC http://mgc.nci.nih.gov/	Unpublished (1999)	Contact: Robert Strausberg, Ph. D. Email: cgabs-remail.nih.gov Tissue Procurement: ATCC/DC/DTP CDNA Library Preparation: Ling Hong/Rubin Laboratory DNA Sequencing by: Inceye Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov plate: L10M1292 row: 1 column: 18 High quality sequence stop: 747.

FEATURES

source

1. 553

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Plate-2134 Col-24 Row-E"

/clone_lib="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: sperm; Vector: pBelBAC11; BAC Clones in E-Coli DH10B"

BASE COUNT

146 a 97 c 105 g 197 t 8 others

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 3.5;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

737 TACACAATGTAGACCAAGAG 758


```

Db      531 TACACAATGTAGACCAAG 510
RESULT  3
LOCUS   CENS0363A
DEFINITION Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
          215173 of library G from Tetraodon nigroviridis, genomic survey
          sequence.
ACCESSION AL228519
VERSION   AL228519.1 GI:7888514
KEYWORDS  GSS; genome survey sequence.
SOURCE    Tetraodon nigroviridis.
ORGANISM  Tetraodon nigroviridis.
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
          Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
          Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 909)
AUTHORS   Roest-Crollius,H., Jallion,O., Dasilva,C., Fzames,C., Fisher,C.,
          Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
          Weissenbach,J.
TITLE     Characterization and repeat analysis of the compact genome of the
          freshwater pufferfish Tetraodon nigroviridis
JOURNAL   Unpublished
AUTHORS   2 (bases 1 to 909)
REFERENCE 2 (bases 1 to 909)
AUTHORS   Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
          Bernot,A., Fzames,C., Wincker,P., Brothier,P., Quetier,F.,
          Saurin,W. and Weissenbach,J.
TITLE     Human gene number estimate provided by genome wide analysis using
          Tetraodon nigroviridis DNA sequence
JOURNAL   Unpublished
AUTHORS   3 (bases 1 to 909)
REFERENCE 3 (bases 1 to 909)
AUTHORS   Genoscope.
TITLE     Direct Submission
COMMENT    Submitted (12-APR-2000) to the EMBL/Genbank/DDJ databases
          This sequence is a single read and was generated as part of a large
          scale clone-end sequencing project of the Tetraodon nigroviridis
          genome. For more information, please take a look at
          http://www.genoscope.cns.fr/Tetraodon.
FEATURES
          source
          1..909
          /organism="Tetraodon nigroviridis"
          /db_xref="taxon:99883"
          /clone="215123"
          /clone_1lb="G"
          /note="Genoscope sequence ID : C0AG215AE12SP1-end :
          PUC-ori"
BASE COUNT 226 a 205 c 243 g 228 t 7 others
ORIGIN
Query Match 1.1%; Score 22; DB 12; Length 909;
Best Local Similarity 100.0%; Pred.No. 3.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1257 CAGACAAAAAATGTGAGCAAG 1278
Db 567 CAGACAAAAAATGTGAGCAAG 588
RESULT 4
LOCUS   AL632585
DEFINITION AL632585 XGC.gastrula silurana tropicalis cDNA clone Tgas021e05 5',
          mRNA sequence.
ACCESSION AL632585
VERSION   AL632585.1 GI:16602068
KEYWORDS  EST.
SOURCE    western clawed frog.
          Silurana tropicalis
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 389)
AUTHORS   Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
TITLE     Sanger Xenopus tropicalis EST project 2001 (10-2001)
JOURNAL   Unpublished (2001)
COMMENT    Contact: Huckle E
          Sanger Centre
          Hinxton, Cambridgeshire, CB10 1SA, UK
          Email: trop@sanger.ac.uk
          Sanger Xenopus tropicalis EST project 2001
          TROPICALIS_SEQUENCE_ID: Tgas021e05.sp6
          Sequencing primer: SP6
          This sequence is from a Xenopus Gene Collection (XGC) library
          constructed by Aaron M. Zorn.
FEATURES
          source
          1..389
          /organism="Silurana tropicalis"
          /db_xref="taxon:8364"
          /clone="Tgas021e05"
          /clone_1lb="XGC-gastrula"
          /dev_stage="gastrula (stages 10.5-13 mixed)"
          /lab_host="Escherichia coli XL1-blue"
          /note="Vector: PCS107; Site 1: EcoRI; Site 2: NotI; cDNA
          was oligo dt primed from 5ug of poly A+ RNA from stages
          10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
          into PCS107 with EcoRI at the 5' end and NotI at the 3'
          end."
BASE COUNT 93 a 110 c 104 g 82 t
ORIGIN
Query Match 1.1%; Score 21; DB 9; Length 389;
Best Local Similarity 100.0%; Pred.No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1429 TATGATGACCGACATGATGAA 1449
Db 312 TATGATGACCGACATGATGAA 332
RESULT 5
LOCUS   AI397462
DEFINITION AI397462 422 bp mRNA linear EST 07-JUN-2001
          FB18P08 xl zebrafish Washu MPIM EST Danio rerio cDNA clone
          IMAGE:37112215 3', similar to gb:X57527 COLLAGEN ALPHA 1(VIII) CHAIN
          PRECURSOR (HUMAN); mRNA sequence.
ACCESSION AI397462
VERSION   AI397462.1 GI:4227355
KEYWORDS  EST.
SOURCE    zebrafish.
          Danio rerio
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
          1 (bases 1 to 422).
REFERENCE 1 (bases 1 to 422)
AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
          S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
          K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
          Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurr,R., Ritter,E.,
          Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
          and Wilson,R.
TITLE     Washu zebrafish EST Project 1998
JOURNAL   Unpublished (1998)
COMMENT    Contact: Stephen L. Johnson
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: zbrafish@watson.wustl.edu
          cDNA Library Preparation: Matthew Clark. cDNA library Arrayed by:
          Matthew Clark. DNA Sequencing by: Washington University Genome
          Sequencing Center Clone distribution: Genome Systems, St. Louis,

```


Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and RessourcenzentrumPrimatendbank, Berlin, Germany (web address: www.tzpd.de)
Seq primer: T7 ET from Amersham
High quality sequence stop: 415
POLYA-No.

FEATURES

SOURCE

Location/Qualifiers
1..422
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone_image="3712215"
/clone_lib="zebrafish Washu MPIMG EST"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield stage embryos"
/lab_host="XLI-blue MRF"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st strand cDNA was primed with a Not I - oligo(dT)15 primer [5,pgACTACTCTAGATCGGAGCGGCCCGCTTTTCTTTTCTT3']; double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach lab, ICRF, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

BASE COUNT 120 a 88 c 97 g 117 t
ORIGIN

Query Match 1.1%; Score 21; DB 9; Length 422;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1757 GGCTCTACAGAAATGGAGA 1777
|||||
Db 236 GGCTCTACAGAAATGGAGA 216

RESULT 6
AL629478 704 bp mRNA linear EST 02-NOV-2001
LOCUS AL629478
DEFINITION AL629478 XGC-gastrula silurana tropicalis cDNA clone Tgas018104 5', mRNA sequence.
ACCESSION AL629478
VERSION AL629478.1 GI:16598961
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Silurana.
1 (bases 1 to 704)
Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (10_2001)
Unpublished (2001)
Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: tropesanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: Tgas018104.sp6
Sequencing primer: SP6

This sequence is from a Xenopus Gene Collection (XGC) library constructed by Aaron M. Zorn.
Location/Qualifiers
1..704
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone_image="Tgas018104"
/clone_lib="XGC-gastrula"
/dev_stage="gastrula (stages 10-5-13 mixed)"
/lab_host="Escherichia coli XLI-blue"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from stages 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."

BASE COUNT 153 a 213 c 188 g 150 t
ORIGIN

Query Match 1.1%; Score 21; DB 9; Length 704;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 TATGATGACGACATGATGAA 1449
|||||
Db 318 TATGATGACGACATGATGAA 338

RESULT 7
CNS06N80 1025 bp DNA linear GSS 17-JUN-2001
LOCUS CNS06N80
DEFINITION T3 end of clone AU00A010D02 of library AU00A from strain CBS 3082 of Saccharomyces kluyveri, genomic survey sequence.
ACCESSION AL406462
VERSION AL406462.1 GI:12170499
KEYWORDS GSS.
SOURCE Saccharomyces kluyveri.
ORGANISM Saccharomyces kluyveri
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE 1 (bases 1 to 1025)
AUTHORS Neuveglise,C., Bon,E., Lepingle,A., Wincker,P., Artiguenave,F., Gaillardin,C. and Casaregola,S.
TITLE Genomic exploration of the hemiascomycetous yeasts: 9. Saccharomyces kluyveri
JOURNAL FEBS Lett. 487 (1), 56-60 (2000)
MEDLINE 20584719

2 (bases 1 to 1025)
Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brothier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potter,S., Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE 20584711

3 (bases 1 to 1025)
Genoscope.
Direct Submission
Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail: seqref@genoscope.cns.fr - Web: www.genoscope.cns.fr)

COMMENT
This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of

the other extremity of this insert.

FEATURES
source
Location/Qualifiers
1..1025

/organism="Saccharomyces kluyveri"
/strain="CBS 3082"
/db_xref="taxon:4934"
/clone="AU00A010D02"
/clone_11b="AU00AA"
/note="end : T3"
<30..>1025

misc-feature
/note="similar to Saccharomyces cerevisiae ORF YGR233c [PH081 : cyclin-dependent kinase inhibitor]"
/evidence=not_experimental

BASE COUNT 348 a 191 c 205 g 280 t 1 others

ORIGIN

Query Match 1.1%; Score 21; DB 12; Length 1025;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1460 ACTATGGAAGTGCAGCAAT 1480
DB 590 ACTATGGAAGTGCAGCAAT 610

RESULT 8
CNS06NM7/C 1078 bp DNA linear GSS 04-JUL-2001

DEFINITION T7 end of clone AU00A015C12 of library AU00A from strain CBS 3082
of Saccharomyces kluyveri, genomic survey sequence.

ACCESSION AL406949
VERSION AL406949.1 GI:12171649

KEYWORDS GSS.
ORGANISM Saccharomyces kluyveri.

REFERENCE
AUTHORS Saccharomyces kluyveri.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

1 (bases 1 to 1078)
Neueglise, C., Bon, E., Lepingle, A., Winker, P., Attienave, F.,
Gallard, C., and Casaregola, S.
Genomic exploration of the hemiascomycetous yeasts: 9.
Saccharomyces kluyveri

FEB5 Lett. 487 (1), 56-60 (2000)

JOURNAL
MEDLINE
REFERENCE
AUTHORS

2 (bases 1 to 1078)

Souci, J. L., Aigle, M., Attienave, F., Blandin, G.,
Boitlin-Fukuhara, M., Bon, E., Brothier, P., Casaregola, S.,
de-Montigny, J., Dujon, B., Durand, P., Lepingle, A., Liorente, B.,
Malpertuy, A., Neueglise, C., Ozier-Kalogeropoulos, O., Potier, S.,
Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,
Winker, P., and Weissenbach, J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies

FEB5 Lett. 487 (1), 3-12 (2000)

3 (bases 1 to 1078)

Genoscope.
Direct Submission
Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
anastus, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.

COMMENT

JOURNAL
MEDLINE
REFERENCE
AUTHORS

Location/Qualifiers

FEATURES

source

1..1078
/organism="Saccharomyces kluyveri"
/strain="CBS 3082"
/db_xref="taxon:4934"
/clone="AU00A015C12"
/clone_11b="AU00AA"
/note="end : T7"
complement(1..>1074)

misc-feature
/note="similar to Saccharomyces cerevisiae ORF YGR233c [PH081 : cyclin-dependent kinase inhibitor]"
/evidence=not_experimental

BASE COUNT 300 a 218 c 214 g 345 t 1 others

ORIGIN

Query Match 1.1%; Score 21; DB 12; Length 1078;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1460 ACTATGGAAGTGCAGCAAT 1480
DB 739 ACTATGGAAGTGCAGCAAT 719

RESULT 9
BF174868

LOCUS BF174868 237 bp mRNA linear EST 23-MAR-2001
DEFINITION MYE3951 Myeloma (MYE) CDNA library Homo sapiens cDNA, mRNA
sequence.

ACCESSION BF174868
VERSION BF174868.1 GI:13441082

KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 237)
Claudio, J. O., Tang, H., Khan, E. M., Vorralla, M., Li, Z., Cukerman, E.,
Francisco-Pabalan, O., Ilew, C. C. and Stewart, A. K.

The transcriptional phenotype of myeloma cells
Unpublished (2000)
Contact: A. Keith Stewart, M.D.
Oncology Research
University Health Network
610 University Ave., 5-126, Toronto, Ontario, M5G 2M9, Canada
Tel: (416) 946-4639
Fax: (416) 946-6546
Email: k.stewart@utoronto.ca

PCR Primers
FORWARD: 5'-GCCAAGCTCGAATTAACCTCACTAAGGCG-3'
BACKWARD: 5'-CCAGTGAATTTGTAATACGACTCACTAAGGCG-3'
Seq primer: 5'-GAATTAACCTCACTAAGG-3'.

Location/Qualifiers
1..237
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_11b="Myeloma (MYE) cDNA library"
/sex="male"
/tissue="Blood"
/cell_type="myeloma"
/dev_stage="multiple myeloma"

FEATURES
source

/note="vector: lambda zap Express; Site1: EcoRI; Site2:
XhoI; Myeloma cells from multiple myeloma patients; bone
marrow were purified by magnetic cell sorting. mRNA were
restriction and an oligo d(T)18 primer containing XhoI
restriction site was used to prime first strand synthesis
using M-MLV reverse transcriptase. To protect the cDNAs
from XhoI digestion in subsequent cloning step, the
nucleotide analogue 5-methyl-dCTP was added to the
nucleotide mixture and [a-32P]dATP was added to monitor
the quantity and quality of first strand synthesis. After
second-strand synthesis and blunting of cDNA termini,
EcoRI adapters were ligated, followed by kinase treatment

and digestion with XhoI. The cDNAs were then size-fractionated using Sephacryl S-500 column and then ligated into EcoRI and XhoI digested Lambda Zap Express vector. The ligation product was packaged using Gigapack II packaging extract. The library was packaged using Gigapack II packaging extract. The library had primary titre of approx. 1x10⁶. Clones from the primary library were randomly selected for single pass sequencing.

BASE COUNT 103 a 35 c 54 g 45 t

Query Match 1.0%; Score 20; DB 10; Length 237;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCAACCAAGCAAGAAA 338
|||||
Db 191 AAAAGCAACCAAGCAAGAAA 210

RESULT 10
LOCUS AA317576 287 bp mRNA linear EST 19-APR-1997
DEFINITION EST19524 Retina II Homo sapiens cDNA 5' end similar to DNA-binding protein PO-GA, mRNA sequence.
ACCESSION AA317576
VERSION AA317576.1 GI:1969905
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 287)
Adams,M.D., Kervatage,A.R., Fleischmann,R.D., Fuldner,R.A., Bolt,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Colton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodok,A., Guelm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M., Kelley,D.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utechtack,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferris,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weisner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT Other-ESTs: TH0171722

Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018689056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/cdb/hg1/hg1.html>)
Seq primer: M13 Reverse.

FEATURES
source location/Qualifiers
1. 287

/organism="Homo sapiens"
/db_xref="ATCC (inhost):118036"
/db_xref="taxon:9606"
/clone_lib="Retina II"
/sex="mixed"

/dev_stage="adult"
/note="Organ: retina; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 113 a 57 c 64 g 52 t 1 others

Query Match 1.0%; Score 20; DB 9; Length 287;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCAACCAAGCAAGAAA 338
|||||
Db 243 AAAAGCAACCAAGCAAGAAA 262

RESULT 11
LOCUS BB551207 295 bp mRNA linear EST 01-AUG-2000
DEFINITION BB551207 RIKEN full-length enriched, 2 days pregnant adult female oviduct Mus musculus cDNA clone E230026H21 3' similar to D10712 Mus musculus mRNA for nedd-1 protein, mRNA sequence.
ACCESSION BB551207
VERSION BB551207.1 GI:9637573
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 295)
Kono,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,Kono,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya,T., Tsunoda,Y., Watabiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Kono,H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki.
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermocyclization of labile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Katsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES
source location/Qualifiers
1. 295

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E230026H21"


```

/oranism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone_image="3471951"
/clone_id="Soares_thymus_2NBM1"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/notes="Vector: pT73D-pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand
cDNA was primed with a Not I - oligo(dT) primer [5',

```

Query Match	1.0%;	Score 20;	DB 10;	Length 367;
Best Local Similarity	100.0%;	Pred. No. 38;		

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 325 AACCAAGCAGAAATCTTGA 344
 Db 242 AACCAAGCAGAAATCTTGA 261

RESULT 14

BM173589 402 bp mRNA linear EST 04-DEC-2001
 LOCUS 900943 Avicennia marina leaf cDNA library Avicennia marina cDNA
 DEFINITION clone AM900943.5' similar to NADH dehydrogenase subunit 5 (AF085145
) of Schistosoma mansoni, mRNA sequence.

ACCESSION BM173589
 VERSION BM173589.1 GI:17313152
 KEYWORDS EST.
 SOURCE Avicennia marina.
 ORGANISM Avicennia marina.

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Lamiales; Avicenniaceae; Avicennia.
 TITLE 1 (bases 1 to 402)
 JOURNAL Parani M., Mehta P., Sivaprakash, K.R. and Parida, A.
 COMMENT Unpublished (2000)
 Expressed sequence tags from the mangrove species Avicennia marina
 Contact: Parani M / Parida A
 Department of Plant Molecular Biology
 M. S. Swaminathan Research Foundation
 111 Cross Street, Taramani Institutional Area, Chennai 600 113,
 India

REFERENCE Parani M., Mehta P., Sivaprakash, K.R. and Parida, A.
 AUTHORS Expressed sequence tags from the mangrove species Avicennia marina
 JOURNAL Unpublished (2000)
 COMMENT Contact: Parani M / Parida A
 Department of Plant Molecular Biology
 M. S. Swaminathan Research Foundation
 111 Cross Street, Taramani Institutional Area, Chennai 600 113,
 India

Tel: 91-44-2351319
 Fax: 91-44-2351319
 Email: mangrovegenes@msrf.res.in
 BLAST search in BLASTX (Non-redundant) using default parameters as
 on November 6th 2001
 Seq primer: M13 Reverse.

FEATURES

source
 1. 402
 Location/Qualifiers
 /organism="Avicennia marina"
 /strain="Pichavaram"
 /db_xref="taxon:82927"
 /clone_11b="Am900943"
 /clone_11b="Avicennia marina leaf cDNA library"
 /sex="Hermaphrodite"
 /tissue_type="leaf"
 /dev_stage="1 year old plant"
 /note="A one-year-old plant from the natural mangrove
 habitat in Pichavaram, Tamil Nadu, India was collected
 and brought to the M.S. Swaminathan Research Foundation.
 The whole plant was treated with the nutrient solution
 supplemented with 500 mM NaCl for 48 hours. Poly(A+) mRNA
 from the leaf tissue was purified and cDNA was prepared
 using Superscript Lambda System (Life Technologies, Cat.
 No. 19643-014). The cDNAs were size fractionated over
 SizeSep-400 spun column (Amersham-Pharmacia Biotech Cat
 No. 27-5105-01) and cloned in 5' Sal I - 3' Not I of
 pSPORT1 (Life Technologies Cat. No. 15383-011). The ESTs
 were sequenced from the 5' end using M13/pUC18 reverse
 primer in an Automated Sequencer (ABI310). Applied
 Biosystems) and submitted after editing to remove the
 vector and adapter sequences. BLAST search in BLASTX
 (Non-redundant) was carried out on 7th August, 2000 using
 default parameters and the results reported under
 putative identification/comment"

BASE COUNT 101 a 77 c 83 g 141 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 402;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1281 GTGAAAGATGCTGCTGCT 1300

Db 341 GTGAAAGATGCTGCTGCT 360

RESULT 15

BF016982 405 bp mRNA linear EST 10-OCT-2000
 LOCUS ux19c09.y1 Soares_thymus_2NDBMT Mus musculus cDNA clone
 DEFINITION IMAGE:3471952.5' similar to SW:NEDI_MOUSE P33215 NEDD1 PROTEIN ;
 mRNA sequence.

ACCESSION BF016982
 VERSION BF016982.1 GI:10748314
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 TITLE 1 (bases 1 to 405)
 JOURNAL NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Other_ESTs: ux19c09.x1
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:1391312

Seq primer: -40RP from Gibco
 High quality sequence stop: 344.

FEATURES

source
 1. 405
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_11b="IMAGE:3471952"
 /clone_11b="Soares_thymus_2NDBMT"
 /sex="male"
 /tissue_type="Thymus"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 3'] : double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through two
 rounds of normalization, and was constructed by Bento
 Soares and M. Fatima Bonaldo."

BASE COUNT 124 a 101 c 88 g 92 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 405;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 325 AACCAAGCAGAAATCTTGA 344
 Db 283 AACCAAGCAGAAATCTTGA 302

RESULT 16

AQ445724 406 bp DNA linear GSS 16-DEC-1999
 LOCUS nbx0054J07f CUGI Rice BAC Library Oryza sativa genomic clone
 DEFINITION nbx0054J07f, DNA sequence.

ACCESSION AQ445724
 VERSION AQ445724.2 GI:6582598
 KEYWORDS GSS.
 SOURCE Oryza sativa.

ORGANISM *Oryza sativa*
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrharioideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 406)
 Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Rice Genome Unpublished (1998)
 On Dec 15, 1999 this sequence version replaced gi:4574947.
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: TRATRCAGCTCCTATGAGG
 Class: BAC ends
 High quality sequence stop: 376.

FEATURES
 source
 1..406
 Location/Qualifiers
 /organism="Oryza sativa"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /clone="nbxb0054j07f"
 /clone_lib="CGI Rice BAC library"
 /tissue_type="leaf"
 /lab_host="E. coli DH10B"
 /note="Vector: pBel0AC11; Site_1: HindIII; Site_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from *Oryza sativa*, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9%. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

BASE COUNT 129 a 77 c 64 g 136 t
 ORIGIN

Query Match 1.0%; Score 20; DB 12; Length 406;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 636 CAAGTAATACCTGCTTAC 655
 ||||||||||||||||||||
 Db 219 CAAGTAATACCTGCTTAC 238

RESULT 17
 BG017796 440 bp mRNA linear EST 30-JAN-2001
 LOCUS debl0f11.y1 NICHD XGC Spl Xenopus laevis cDNA clone IMAGE:4174220
 DEFINITION 5, similar to TR-090Y16 090Y16 MDJ7. ; mRNA sequence.
 ACCESSION BG017796 GI:12472380
 VERSION BG017796.1
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM *Xenopus laevis*
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

REFERENCE 1 (bases 1 to 440)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
 cdNA Library Preparation: Life Technologies, Inc.
 cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LNL at: info@image.lnl.gov
 Seq primer: -40RP from Gibco.

FEATURES
 source
 1..440
 Location/Qualifiers
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="IMAGE:4174220"
 /clone_lib="NICHD XGC Spl"
 /dev_stage="adult"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.4 kb. Constructed by Life Technologies."

BASE COUNT 136 a 83 c 99 g 122 t
 ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 440;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1306 TGGCTTTGCAATGTGGAA 1325
 ||||||||||||||||||||
 Db 335 TGGCTTTGCAATGTGGAA 354

RESULT 18
 BB824552 442 bp mRNA linear EST 19-NOV-2001
 LOCUS BB824552
 DEFINITION BB824552 RIKEN full-length enriched, mammary gland RCB-0526 Jyg-MC(A) cdna Mus musculus cDNA clone G830033120 3', mRNA sequence.
 ACCESSION BB824552 GI:17002795
 VERSION BB824552
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM *Mus musculus*
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 442)
 Akimura,T., Arahata,K., Hirakawa,T., Hirazane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watanuki,A., Yasunishi,A., Yamatsus,M. and Hayashizaki,Y.
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)
 Unpublished (2001)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp,

URL: <http://genome.gsc.riken.go.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 Waga, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
 e mouse tissues.

FEATURES

source location/Qualifiers
 1..442
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="g830033i20"
 /clone_lib="RIKEN full-length enriched, mammary gland
 RCB-0526 Jyg-MC(A) cDNA"
 /tissue_type="mammary gland"
 /cell_line="RCB-0526 Jyg-MC(A)"
 BASE COUNT 138 a 93 c 109 g 102 t
 ORIGIN

Query Match 1.0%; Score 20; DB 9; Length 442;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 325 ACCACAGCAAGAAATCTTGA 344
 ||||||||||||||||
 Db 408 AACCAACGACAGAAATCTTGA 427

RESULT 19
 BE566388 443 bp mRNA linear EST 15-AUG-2000
 LOCUS 601339984F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3682221 5',
 DEFINITION mRNA sequence.
 ACCESSION BE566388
 VERSION BE566388.1 GI:9810108
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 443)
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: L1CM365 row: g column: 22
 High quality sequence stop: 441.
 Location/Qualifiers
 1..443
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

FEATURES

source
 1..443
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

/clone="IMAGE:3682221"
 /clone_lib="NIH_MGC_53"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: bladder; Vector: pDNR-LIB (Clontech);
 Site_1: SfiI (ggcgcttcggcc); Site_2: SfiI (ggccatcatggcc
); Double-stranded cDNA was prepared from cell line RNA.
 5' and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGCCATATTCGCC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCCGCGCGCCGACATG-dT(30)BN-3'
 (where B = A, C, or G and N = A, C, G, or T). Average
 insert size 1.55 kb (range 0.9-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA)."
 BASE COUNT 167 a 91 c 100 g 84 t 1 others
 ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 443;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 319 AAAGCAACCAAGCAAGAAA 338
 ||||||||||||||||
 Db 274 AAAGCAACCAAGCAAGAAA 293

RESULT 20
 BG942702 446 bp mRNA linear EST 11-JUN-2001
 LOCUS ax28e04.x1 Proliferating Human Erythroid Cells (LCB:ax library)
 DEFINITION Homo sapiens cDNA clone ax28e04 random, mRNA sequence.
 ACCESSION BG942702
 VERSION BG942702.1 GI:14342074
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 446)
 REFERENCE Gubln, A.N., Njoroge, J.M., Bouffard, G.G. and Miller, J.L.
 TITLE Gene expression in proliferating human erythroid cells
 JOURNAL Genomics 59 (2), 168-177 (1999)
 MEDLINE 99339981
 COMMENT Contact: Jeffery L. Miller
 Laboratory of Chemical Biology
 National Institute of Diabetes and Digestive and Kidney Diseases
 Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
 20892, USA
 Tel: 301 402 2373
 Fax: 301 435 5148
 Email: jm7fe@nih.gov
 DNA Sequencing and analyses by National Institutes of Health
 Intramural Sequencing Center (NISC).
 Plate: 28 row: e column: 04
 Seq primer: -21M13 forward primer (ABI).
 Location/Qualifiers
 1..446
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ax28e04"
 /clone_lib="Proliferating Human Erythroid Cells (LCB:ax
 library)"
 /sex="unknown"
 /tissue_type="blood"
 /cell_line="Erythroid Cells"
 /cell_line="Primary Culture of Peripheral Blood
 Mononuclear Cells"
 /dev_stage="Progenitor; EPO responsive CD11++++"
 /lab_host="SOLR"
 /note="Organ: Blood; Vector: Lambda ZAP II; Site_1: EcoRI;
 Site_2: EcoRI; 65,000 proliferating erythroid cells from

the buffy coat of a blood donation were obtained by flow cytometric separation after a 5-day culture period in the presence of erythropoietin. Total RNA was purified from the sorted cell population using TRIzol reagent. RNA (0.3 ug) was converted into double stranded cDNA using Clontech's Capfinder cDNA Library Construction Kit (Clontech) according to the manufacturer's protocol and cloned into EcoRI digested lambda zap II vector (Stratagene). The phage library was amplified once prior to in vivo excision in SOLR cells. Individual colonies were grown, and the cDNA inserts were sequenced in high throughput (NIH Intramural sequencing center <http://www.nisc.nih.gov/>).

BASE COUNT 171 a 85 c 102 g 88 t
ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 446;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAGCAGCAAGCAAGAA 338
DB 260 AAAGCAGCAAGCAAGAA 279

RESULT 21
LOCUS BB750480 450 bp mRNA linear EST 16-OCT-2001

DEFINITION BB750480 RIKEN full-length enriched, pooled tissues, cerebellum, etc. Mus musculus cDNA clone G130003M12.3', mRNA sequence.

ACCESSION BB750480
VERSION BB750480.1 GI:16154716

KEYWORDS EST.
SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shireki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Wataniki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length CDNAs (Akimura,T., et al. 2001)

Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Sucho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gs.c.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

Waki,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichannel sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

FEATURES
source location/Qualifiers

1. 450
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G130003M12"
/clone_lib="RIKEN full-length enriched, pooled tissues, cerebellum, etc."
/note="pooled tissues : (tissue_type=cerebellum, dev_stage=16 days neonate, sex=mixed), (tissue_type=cerebellum, dev_stage=0 day neonate, sex=mixed), (tissue_type=hippocampus, dev_stage=adult, sex=male), (tissue_type=whole body, dev_stage=9 days embryo, sex=mixed), (tissue_type=lung, dev_stage=13 days embryo, sex=mixed)"
BASE COUNT 141 a 94 c 106 g 109 t
ORIGIN

Query Match 1.0%; Score 20; DB 9; Length 450;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 325 AACCAAGCAAGAAATCTTGA 344
DB 416 AACCAAGCAAGAAATCTTGA 435

RESULT 22

LOCUS AA280963

DEFINITION AA280963 493 bp mRNA linear EST 15-AUG-1997
zs97d12.r1 NCI-CGAP GCBI Homo sapiens cDNA clone IMAGE:711575 5' similar to gb:U23320 ACTIVATOR 1 140 KD SUBUNIT (HUMAN);, mRNA sequence.

ACCESSION AA280963
VERSION AA280963.1 GI:1923644

KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
1 (bases 1 to 493)

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.jnl.gov) for further information.

Seq primer: 28m13 rev2 ET from Amersham
High quality sequence stop: 469.

location/Qualifiers
1. 493

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:711575"
/clone_lib="NCI-CGAP GCBI"
/tissue_type="germlinal center B cell"
/lab_host="DH10B"

/note="Vector: pT73D-Pac (pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Altman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer

primed with a Not I - oligo(dT) primer

[5'-TGTACCAATCTGAAGTGGAGCGCCGCTCATTTTCTTTTCTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pRT3 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 188 a 94 c 115 g 96 t

Query Match 1.0%; Score 20; DB 9; Length 493;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCAACCAAGCAAGAA 338
 Db 274 AAAAGCAACCAAGCAAGAA 293

RESULT 23
 LOCUS AW762116 526 bp mRNA linear EST 04-MAY-2000
 DEFINITION ur53h07.y1 NCI-CGAP_Mam3 Mus musculus cDNA clone IMAGE:3154045 5'
 (MOUSE); mRNA sequence.

ACCESSION AW762116 GI:7694034
 VERSION AW762116
 KEYWORDS EST.

SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 526)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index

JOURNAL Unpublished (1997)
 COMMENT Other-ESTs: ur53h07.x1
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
image.lnl.gov/image/html/iresources.shtml

MG1:1056801
 Seq primer: -40RP from Gibco
 High quality sequence stop: 405.
 Location/Qualifiers

1..526
 /organism="Mus musculus"
 /strain="129 - C57/B6 - FVBN"
 /db_xref="taxon:10090"
 /clone="IMAGE:3154045"
 /clone_lib="NCI-CGAP Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="DH10B"

/note="Organ: mammary; Vector: PCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
 Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999)."

BASE COUNT 159 a 130 c 123 g 114 t

Query Match 1.0%; Score 20; DB 9; Length 526;
 Best Local Similarity 100.0%; Pred. No. 41;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 325 AACCAAGCAAGAACTTGA 344
 Db 299 AACCAAGCAAGAACTTGA 318

RESULT 24
 LOCUS AO710693 526 bp DNA linear GSS 13-JUL-1999
 DEFINITION HS-5344_A2_E10_T7A RCI-11 Human Male BAC Library Homo sapiens
 genomic clone Plate=920 Col=20 Row=I, DNA sequence.
 ACCESSION AO710693
 VERSION AO710693.1 GI:5460009
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 526)

Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome

Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
 99380589

CONTACT: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887

Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RCI-11. For BAC
 library availability, please contact Pletier de Jong
 (pletierdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
 or from Research Genetics (info@resgen.com). BAC end Web Server:
<http://www.htsc.washington.edu>

Plate: 920 row: I column: 20
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 526.
 Location/Qualifiers

1..526
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate=920 Col=20 Row=I"
 /clone_lib="RCI-11 Human Male BAC Library"
 /sex="male"

/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
 male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBACe3.6 vector at EcoRI sites"

BASE COUNT 123 a 103 c 113 g 185 t 2 others

ORIGIN

Query Match 1.0%; Score 20; DB 12; Length 526;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1699 CTCATGATGACCTACTGT 1718
 Db 431 CTCATGATGACCTACTGT 450

RESULT 25
 LOCUS AZ824983 535 bp DNA linear GSS 20-FEB-2001
 DEFINITION 2M0099J24R Mouse 10kb plasmid UUCIM library Mus musculus genomic

clone UUGC2M009J24 R, DNA sequence.
 ACCESSION A2824983
 VERSION A2824983.1 GI:12994891
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 535)
 REFERENCE Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A. and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 TITLE Unpublished (2000)
 JOURNAL Contact: Robert B. Weiss
 COMMENT University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0099 row: J column: 24
 Seq primer: CACACGAGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 535.
 Location/Qualifiers
 1..535
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M009J24"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (9147321141gblAF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 164 a 130 c 120 g 121 t
 ORIGIN

Query Match 1.0%; Score 20; DB 12; Length 535;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 325 AACCAAGCAAGAAATCTTGA 344
 ||||||||||||||||||
 DB 301 AACCAAGCAAGAAATCTTGA 320

RESULT 26
 BG723460 563 bp mRNA linear EST 08-MAY-2001
 LOCUS

DEFINITION 602694145F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4826218 5', mRNA sequence.
 ACCESSION BG723460
 VERSION BG723460.1 GI:14002647
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 563)
 REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaabs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA library preparation: Michael J. Brownstein (NHGRI), Shitaki Toshitsuki and Piero Carninci (RIKEN)
 CDNA library arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1AM10740 row: J column: 11
 High quality sequence stop: 406.
 Location/Qualifiers
 1..563
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4826218"
 /clone_lib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescript (modified pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.2 kb and normalized to 800 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NHGRI/NHGRI, National Institutes of Health). Note: this is a NIH_MGC library."

BASE COUNT 201 a 124 c 129 g 109 t
 ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 563;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAGCAACCAAGCAAGAAA 338
 ||||||||||||||||||
 DB 287 AAAGCAACCAAGCAAGAAA 306

RESULT 27
 BG503218 587 bp mRNA linear EST 27-MAR-2001
 LOCUS 602551318F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4658644 5', mRNA sequence.
 ACCESSION BG503218
 VERSION BG503218.1 GI:13464735
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 587)
 REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>

Plate: LNCM1453 row: d column: 05
 High quality sequence stop: 587.

FEATURES
 source
 1. 587
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4658644"
 /clone_1lb="NIH_MGC_61"
 /tissue_type="embryonal carcinoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
 Still (ggcgccctggcc); Site_2: Still (ggccattatggcc);
 Double-stranded cDNA was prepared from cell line RNA. 5'
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-ATTCTAGAGCGGCGCCGACATG-dt(30)-BN-3'
 (where B = A, C, or G and N = A, C, G, or T). Average
 insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
 Library."
 BASE COUNT 237 a 112 c 126 g 112 t
 ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 587;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAAGCAACGACGAGAAA 338
 ||||||||||||||||||
 Db 268 AAAAGCAACGACGAGAAA 287

RESULT 28 588 bp mRNA linear EST 03-APR-2001
 BG539230 602567826F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4692414 5',
 LOCUS mRNA sequence.
 DEFINITION
 ACCESSION BG539230
 VERSION BG539230.1 GI:13531463
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 588)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LNCM1513 row: c column: 07
 High quality sequence stop: 586.

FEATURES
 source
 1. 588
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4692414"

/clone_1lb="NIH_MGC_77"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
 Still (ggcgccctggcc); Site_2: Still (ggccattatggcc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCGGCGCCGACATG-dt(30)-BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.9
 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."
 BASE COUNT 227 a 113 c 130 g 118 t
 ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 588;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAAGCAACGACGAGAAA 338
 ||||||||||||||||||
 Db 289 AAAAGCAACGACGAGAAA 308

RESULT 29 642 bp mRNA linear EST 30-SEP-1997
 AA606845
 LOCUS
 DEFINITION
 vm94h07.r1 Knowles Solter mouse blastocyst B1 mus musculus cDNA 1
 clone IMAGE:1005949 5' similar to gb:DI0712 Mouse mRNA for nebd-1
 protein, complete cds (MOUSE);, mRNA sequence.
 ACCESSION AA606845
 VERSION AA606845.1 GI:2455738
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 642)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Schellenger,K., Stepcoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:570165

High quality sequence stop: 473.

FEATURES
 source
 1. 642
 Location/Qualifiers

/organism="Mus musculus"
 /strain="B6D2 F1/J"
 /db_xref="taxon:10090"
 /clone="IMAGE:1005949"
 /clone_1lb="Knowles Solter mouse blastocyst B1"
 /tissue_type="blastocyst"
 /dev_stage="embryo (pre-implantation)"
 /lab_host="DH10B"
 /note="Organ: embryo; Vector: pSPORT; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally from mRNA prepared
 from 800 blastocysts. Primer: Salt(dT):
 5'-CGGTGACCGTCGACCGTTTCTTTTCTT-3'. cDNAs were
 cloned into the NotI/SalI sites of a pSPORT vector (Life
 Technologies). Two different size selections: B1 (larger

BASE COUNT 196 a 138 c 143 g 165 t
 ORIGIN

Query Match 1.0%; Score 20; DB 9; Length 642;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 325 AACCAAGCAAGAAATCTGA 344
 DB 74 AACCAAGCAAGAAATCTGA 93

RESULT 30
 LOCUS BG492386 650 bp mRNA linear EST 27-MAR-2001
 DEFINITION 602536313P1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4655307 5',
 mRNA sequence.
 ACCESSION BG492386
 VERSION BG492386.1 GI:13453898
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 650)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LML at:
<http://image.llnl.gov>
 Plate: LCM1444 row: 1 column: 04
 High quality sequence stop: 649.
 Location/Qualifiers

FEATURES
 Source
 1. 650
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4655307"
 /clone_lib="NIH_MGC_59"
 /tissue_type="mucoepidermoid carcinoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
 SfiI (ggcgccctcgcc); Site_2: SfiI (ggcattatggcc);
 Double-stranded cDNA was prepared from cell line RNA. 5'
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGGCCATATGGCC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3'
 (where B = A, C, or G and N = A, C, G, or T). Average
 insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
 Library."

BASE COUNT 258 a 121 c 141 g 129 t 1 others
 ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 650;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAAGCAACGAAGAA 338
 DB 274 AAAAGCAACGAAGAA 293

RESULT 31
 LOCUS BB498524 661 bp mRNA linear EST 25-OCT-2001
 DEFINITION BB498524 RIKEN full-length enriched, 0 day neonate kidney Mus
 musculus cDNA clone D630015B03 3', mRNA sequence.
 ACCESSION BB498524
 VERSION BB498524.2 GI:16442092
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 661)
 AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
 Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,
 M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,
 D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
 Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
 Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 Unpublished (2001)
 On Jul 26, 2000 this sequence version replaced gi:9473625.
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/,
Carninci.P., Shibata.Y., Hayatsu.N., Sugahara.Y., Shibata.K., Itoh
 M., Konno.H., Okazaki.Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Iwawa,M., Ohara,E.,
 Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
 S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and
 Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
 Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanae,I., Aizawa
 K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
 Hayashizaki,Y.
 Computational Analysis of full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
 further details.
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Location/Qualifiers

FEATURES
 Source
 1. 661
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="D630015B03"
 /clone_lib="RIKEN full-length enriched, 0 day neonate
 kidney"
 /tissue_type="kidney"
 /dev_stage="0 day neonate"
 /lab_host="DH10B"
 /note="Site_1: Sali; Site_2: BamHI; cDNA library was

	Query Match	1.0%;	Score 20;	DB 9;	Length 661;
	Best Local Similarity	100.0%;	Pred. No. 44'		
	Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	325 AACCAAGCAAGAAATCTGA	344			
Db	627 AACCAAGCAAGAAATCTTGA	646			

RESULT	32				
BF212414					
LOCUS	BF212414	669 bp	mRNA	linear	EST 06-NOV-2000
DEFINITION	601813376F1 NIH_MGC_54	Homo sapiens	CDNA clone	IMAGE:4047787 5'	
	mRNA sequence.				
ACCESSION	BF212414				
VERSION	BF212414.1	GI:11105913			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				

REFERENCE	1 (bases 1 to 669)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Plate: LLCM877 row: g column: 20
High quality sequence stop: 633.

```

FEATURES
SOURCE
Location/Qualifiers
1..669
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4047787"
/clone_id="NH_MGC_54"
/tissue_type="from chronic myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
Site_1: Sfil (ggcgctctggcgc); Site_2: Sfil (ggcattatggc
); Double-stranded cDNA was prepared from cell l1ne RNA.
5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCGCAATTATGGC-3' and 3' adaptor
sequence: 5'-ATTCCTAGAGCGCGCGCGCATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."

```

Query Match Similarity 1.0%; Score 20; DB 10; Length 669;
Best Local Similarity 100.0%; Pred. NO. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	319	AAAAGCAACCAAGCAAGAA	338
Db	278	AAAAGCAACCAAGCAAGAA	297

[illegible]

ORGANISM	<i>Homo sapiens</i>
	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi,
	Mammalia: Eutheria: Primates: Catarrhini, Homiidae: Homo.
REFERENCE	1 (bases 1 to 677)
AUTHORS	Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J.

TITLE	Identification and initial characterization of 5000 expressed sequenced tags (ESTs) each from adult human normal and osteoarthritic cartilage cDNA libraries
JOURNAL	Osteoarthritis Cartilage 9 (7), 641-653 (2001)
MEDLINE	21482651
COMMENT	Contact: Sanjay Kumar

GlaxoSmithKline
 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
 Tel: 610-270-7245
 Fax: 610-270-5598
 Email: sanjay_kumar-1@gsk.com
 Seq primer: T7
 Location/Qualifiers
 1..677

FEATURES	source
Location/Qualifiers	1. .677
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone_lib="HOA (Human Osteoarthritic Cartilage)"	
/rseq_type="Cartilage"	
/lab_host="E.coli DH10 B"	
/note="Vector: pSFORT 1; Site_1: SalI; Site_2: NotI; Directional"	
BASE COUNT	253 a 141 c 148 g 133 t 2 others
ORIGIN	

	Query Match	Similarity	100.0%	Score 20:	DB 10:	Length 677:
	Best Local	Similarity	100.0%	Pred. No. 44:		
	Matches	20:	Conservative	0:	Mismatches	0:
OY	319	AAAAGCAACCAAGCAAGAAA	338			
Db	350	AAAAGCAACCAAGCAAGAAA	369			

RESULT	34
AA152866	
LOCUS	705 bp
DEFINITION	m556c03.1 Soares.thymus.2Nbdt Mus musculus cDNA clone IMAGE:582722
ACCESSION	5', similar to gb:DI0712 Mouse mRNA for nedd-1 protein, complete cds
VERSION	(MUSSE):' mRNA sequence.
KEYWORDS	AA152866
	GI:1724534
	EST. 18-FEB-1997

SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 705)
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubucque,T., Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Thelsing,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE	The MashU-HMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	Contact: Marra M/Mouse Est Project WashU-HMI Mouse Est Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (infoimage.llnl.gov) for further information. MG1:357372 Seq primer: -28M13 rev2 from Amersham High quality sequence stop: 466. Location/Qualifiers 1..705 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone_image="582724" /clone_lib="Soares_thymus_2nbwt" /sex="male" /tissue_type="Thymus" /dev_stage="4 weeks" /lab_host="DH10B" /note="(vector: pT7T3-Pac (Pharmacia) with a modified polylinker; site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGGAAGTCGGCGCCGCTTTTTTTTTTTTTTTTTTTTTT 3']"; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT	215 a 163 c 171 g 154 t 2 others
ORIGIN	
Query Match	1.0%; Score 20; DB 9; Length 705;
Best Local Similarity	100.0%; Prid. No. 44;
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	325 AACCAAGCAGAATACTTGA 344 Db 293 AACCAAGCAGAATACTTGA 312
RESULT 35	
LOCUS	Bf658632
DEFINITION	Bf658632 710 bp mRNA linear EST 20-DEC-2000 ma98H01.Y1 Soares.thymus.2NBWT Mus musculus cDNA clone IMAGE:3824881.5' similar to SW:NEO1_MOUSE_P33215 NEDD1 PROTEIN ;, mRNA sequence.
ACCESSION	Bf658632
VERSION	GI:11923766
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 710) NCI-Cgap http://www.ncbi.nlm.nih.gov/ncicgap .
AUTHORS	

FEATURES	source
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT	Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium (infoimage.lnl.gov) for further information. MGI:1461233 Seq primer: -40RP from GIBCO High quality sequence stop: 462. Location/Qualifiers 1. 710 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone_image="3824881" /clone_lib="Soares_thymus_2NBMT" /sex="male" /tissue_type="Thymus" /dev_stage="4 weeks" /lab_host="DH10B" /note="Vector: pT7T3-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTCACCAATCTGAAGTGGAGCGCGCGTGTGTGTGTGTGTGTGTGTGT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT	221 a 157 c 177 g 155 t
ORIGIN	
Query Match	1.0%; Score 20; DB 10; Length 710;
Best Local Similarity	100.0%; Prid. NO. 44;
Matches 20; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	325 AACCAAGCAGAAATCTTGA 344
Db	647 AACCAAGCAGAAATCTTGA 666
RESULT 36	
AG099717/c	712 bp DNA linear GSS 03-NOV-2001
LOCUS	Pan troglodytes DNA, clone: PTB-102B07.F, genomic survey sequence.
DEFINITION	AG099717
ACCESSION	AG099717.1 GI:16720234
VERSION	GSS: GSS (genome survey sequence).
KEYWORDS	Pan troglodytes male lymphoblast DNA, clone_lib: PTB Chimpanzee Male BAC library clone: PTB-102B07.F.
SOURCE	Pan troglodytes
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan. 1 (sites) Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totohi, Y., Watanabe, H. and Sakaki, Y. BAC end sequences of library PTB Unpublished 2 (bases 1 to 712) Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totohi, Y., Watanabe, H. and Sakaki, Y. Direct Submission Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsunmi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpansegsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170) Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
AUTHORS	
COMMENT	

PRIMERS

Sequencing: -21M13

LIBRARY

Vector : PKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1. .712

/organism="Pan troglodytes"

/db_xref="taxon:9598"

/clone="PTB-102807.F"

/sex="male"

/cell_type="lymphoblast"

/clone_lib="PTB Chimpanzee Male BAC Library"

BASE COUNT 157 a 200 c 158 g 192 t 5 others

ORIGIN

Query Match

1.0%; Score 20; DB 12; Length 712;

Best Local Similarity 100.0%; Pred. No. 44;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 453 CCTTGCTGCTGGAATGCT 472

Db 230 CCTTGCTGCTGGAATGCT 211

RESULT 37

BC506942/c

LOCUS

DEFINITION 731 bp mRNA linear EST 27-MAR-2001

601861569F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4071257 5',

ACCESSION

BC506942

VERSION

BC506942.1 GI:13468459

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 731)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: c9apbs-remail.nih.gov

Tissue Procurement: CLONTECH Laboratories, Inc.

cDNA Library Preparation: CLONTECH Laboratories, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: LICM916 row: 1 column: 18

High quality sequence stop: 712.

Location/Qualifiers

1. .731

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4071257"

/clone_lib="NIH_MGC_77"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site_1:

SfiI (ggcgctcgccg); Site_2: SfiI (ggcctatggcc); 5' and

3' adaptors were used in cloning as follows: 5' adaptor

sequence: 5'-CAAGGCGCATATGATG-3' and 3' adaptor sequence:

5'-ATTCTAGAGCGCGAGCGCGGACATG-dT(30)BN-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size 1.9

kb (range 0.5-4.0 kb). 12/15 colonies contained inserts

by PCR. This library was enriched for full-length clones

and was constructed by Clontech Laboratories (Palo Alto,

CA). Note: this is a NIH_MGC Library."

BASE COUNT 190 a 190 c 133 g 218 t

Query Match 1.0%; Score 20; DB 10; Length 731;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1439 GACATGATGAAAACAGTCTT 1458

Db 339 GACATGATGAAAACAGTCTT 320

RESULT 38

AZ354492

LOCUS

DEFINITION 737 bp DNA linear GSS 02-OCT-2000

IM0093115R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC1M0093115 R, DNA sequence.

ACCESSION

AZ354492

VERSION

AZ354492.1 GI:10465987

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

BASE COUNT 231 a 178 c 146 g 181 t 1 others

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 737)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert length: 10000 Std Error: 0.00

Plate: 0093 row: 1 column: 15

Seq primer: CACACAGAGAAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 737.

Location/Qualifiers

1. .737

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0093115"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: pMD22uv, Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pMD42 (g114732114|g1AF12072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

ORIGIN

Query Match

1.0%; Score 20; DB 12; Length 737;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1501 TCAGCGTGTAAACAGCAA 1520

DB 72 TCAGCGTGTAAACAGCAA 91

RESULT 39

LOCUS

BF239667 787 bp mRNA linear EST 14-NOV-2000
DEFINITION 601906930F1 NIH_MGC_54 Homo sapiens CDNA clone IMAGE:4134678 5',

ACCESSION

BF239667

VERSION

BF239667.1 GI:11153589

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 787)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

FEATURES

Location/Qualifiers

source

1..787

FEATURES

Location/Qualifiers

source

1..787

FEATURES

Location/Qualifiers

source

1..787

FEATURES

Location/Qualifiers

source

1..787

FEATURES

Location/Qualifiers

source

1..787

FEATURES

Location/Qualifiers

source

1..787

FEATURES

Location/Qualifiers

source

1..787

FEATURES

Location/Qualifiers

source

1..787

FEATURES

Location/Qualifiers

source

1..787

FEATURES

Location/Qualifiers

source

1..787

ORIGIN

Query Match

1.0%; Score 20; DB 10; Length 788;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAGCAACCAAGCAAGAA 338

DB 242 AAAGCAACCAAGCAAGAA 261

RESULT 41

LOCUS

B1690754 803 bp mRNA linear EST 18-SEP-2001
DEFINITION 603314371F1 NCI_CGAP_Mam6 Mus musculus CDNA clone IMAGE:534320 5',

ACCESSION

B1690754

VERSION

B1690754.1 GI:15653383

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 803)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

FEATURES

Location/Qualifiers

source

1..788

FEATURES

Location/Qualifiers

source

1..788

FEATURES

Location/Qualifiers

source

1..788

FEATURES

Location/Qualifiers

source

1..788

FEATURES

Location/Qualifiers

source

1..788

FEATURES

Location/Qualifiers

source

1..788

FEATURES

Location/Qualifiers

source

1..788

FEATURES

Location/Qualifiers

source

1..788

FEATURES

Location/Qualifiers

source

1..788

FEATURES

Location/Qualifiers

source

1..788

found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 plate: LLAM1900 row: n column: 17
 High quality sequence stop: 749.

FEATURES

source

1. .803
 Location/Qualifiers

/organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_lib="IMAGE:5354320"
 /clone_lib="NCL_CGAP_Mam6"
 /sex="female, virgin"
 /tissue_type="Infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt Site_2: Not; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT

252 a 174 c 182 g 195 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 803;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 325 AACCAACGACAGAACTCTGA 344

Db 150 AACCAACGACAGAACTCTGA 169

RESULT 42

BF667800 841 bp mRNA linear EST 21-DEC-2000
 LOCUS 602122620F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4279709 5',
 DEFINITION mRNA sequence.

ACCESSION BF667800
 VERSION BF667800.1 GI:11941695
 KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 841)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CDNA Library Preparation: CLONTECH Laboratories, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM105 row: g column: 06
 High quality sequence stop: 534.

FEATURES

source

1. .841
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:4279709"
 /clone_lib="NIH_MGC_56"
 /tissue_type="Primitive neuroectoderm"

/lab_host="DH10B (r1 phage-resistant)"
 /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
 SfiI (ggcgctcgcc); Site_2: SfiI (ggcgctcgcc);
 Double-stranded cDNA was prepared from cell line RNA. 5'
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGGCCATTTGGCC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCGCGGCGCCGACATG-DT(30)BN-3'

(where B = A, C, or G and N = A, C, G, or T). Average
 insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA).

BASE COUNT 324 a 136 c 220 g 160 t 1 others

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 841;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAGCAACCAAGCAAGAA 338

Db 81 AAAGCAACCAAGCAAGAA 100

RESULT 43

BE383813 861 bp mRNA linear EST 21-JUL-2000
 LOCUS 601298315F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3628587 5',
 DEFINITION mRNA sequence.

ACCESSION BE383813
 VERSION BE383813.1 GI:9329088
 KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 861)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM13 row: m column: 04
 High quality sequence start: 68
 High quality sequence stop: 558.

FEATURES

source

1. .861
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:3628587"
 /clone_lib="NIH_MGC_19"
 /tissue_type="neuroblastoma"

/lab_host="DH10B (phage-resistant)"
 /note="Organ: brain; Vector: pOT7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dt priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

BASE COUNT 236 a 238 c 214 g 173 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 861;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 74 GATATCCACAGATTGGT 93

Db 559 GATATCCACAGATTGGT 578

RESULT 44
BH466047 873 bp DNA Linear GSS 13-DEC-2001
LOCUS BGVSS80TR BOGV Brassica oleracea genomic clone BGVSS80, DNA
DEFINITION
SEQUENCE
ACCESSION BH466047 GI:17665680
VERSION
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 873)
AUTHORS Town, C.D., Van Aken, S., Uterback, T. and Fraser, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other_GSSs: BGVSS80TF
Contact: Chris Town
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
Source
Location/Qualifiers
1..873
/organism="Brassica oleracea"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BGVSS80"
/clone_lib="BOGV"
/note="Vector: PHOS1, site_1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 259 a 205 c 207 g 202 t
ORIGIN

Query Match 1.0%; Score 20; DB 12; Length 873;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1643 AACCTGTGACACCACTTGT 1662
|||||
Db 186 AACCTGTGACACCACTTGT 205

RESULT 45
BF670847 888 bp mRNA linear EST 21-DEC-2000
LOCUS 60214992991 NIH_MGC_81 Homo sapiens CDNA clone IMAGE:4291278 5',
DEFINITION
RNA sequence.
ACCESSION BF670847
VERSION
KEYWORDS BF670847.1 GI:11944742
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 888)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1135 row: 1 column: 07
High quality sequence stop: 622.

FEATURES
Source
Location/Qualifiers
1..888
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4291278"
/clone_lib="NIH_MGC_81"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: muscle (skeletal); Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgccctcgcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGCGCATTTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCCGAGGCGCCGACATG-dT(30)-BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 308 a 191 c 197 g 192 t
ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 888;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAACGACCAACGCAAGANA 338
|||||
Db 302 AAACGACCAACGCAAGANA 321

RESULT 46
BF696798 896 bp mRNA linear EST 22-DEC-2000
LOCUS 602125331P1 NIH_MGC_56 Homo sapiens CDNA clone IMAGE:4282542 5',
DEFINITION
RNA sequence.
ACCESSION BF696798
VERSION BF696798.1 GI:11982206
KEYWORDS
SOURCE EST.
ORGANISM human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1112 row: m column: 07
High quality sequence stop: 563.

FEATURES
Source
Location/Qualifiers
1..896
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4282542"
/clone_lib="NIH_MGC_56"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgccctcgcc); Site_2: SfiI (ggccattatggcc); Double-stranded CDNA was prepared from cell line RNA. 5'

and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 325 a 185 c 207 g 177 t 2 others

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 896;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCAACCAAGCAGAGAA 338
Db 253 AAAAGCAACCAAGCAGAGAA 272

RESULT 47
LOCUS BF247101 904 bp mRNA linear EST 14-NOV-2000
DEFINITION 60185454F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4074298 5',
mRNA sequence.
ACCESSION BF247101
VERSION BF247101.1 GI:11162177
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 904)
NIH-MGC http://mhc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM924 row: h column: 11
High quality sequence stop: 556.
Location/Qualifiers

FEATURES
source
1. 904
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4074298"
/clone_1ib="NIH_MGC_57"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgctcgccg); Site_2: SfiI (ggcgctatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor
sequence: 5'-ATCTAGAGCGCGCGCGCAGCATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.55 kb (range 0.9-4.0 kb). 12/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."

BASE COUNT 320 a 204 c 208 g 172 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 904;
Best Local Similarity 100.0%; Pred. No. 47;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCAACCAAGCAGAGAA 338
Db 257 AAAAGCAACCAAGCAGAGAA 276

RESULT 48
LOCUS BF697265 917 bp mRNA linear EST 22-DEC-2000
DEFINITION 602129793F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:428686 5',
mRNA sequence.
ACCESSION BF697265
VERSION BF697265.1 GI:11982673
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 917)
NIH-MGC http://mhc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM123 row: i column: 23
High quality sequence stop: 613.
Location/Qualifiers

FEATURES
source
1. 917
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:428686"
/clone_1ib="NIH_MGC_56"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgctcgccg); Site_2: SfiI (ggcgctatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor
sequence: 5'-ATCTAGAGCGCGCGCGCAGCATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."

BASE COUNT 343 a 187 c 229 g 158 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 917;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCAACCAAGCAGAGAA 338
Db 272 AAAAGCAACCAAGCAGAGAA 291

RESULT 49
LOCUS BG282996 938 bp mRNA linear EST 21-FEB-2001
DEFINITION 602406185F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4518260 5',
mRNA sequence.
ACCESSION BG282996
VERSION BG282996.1 GI:13032433

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 938)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LHAM10412 row: 3 column: 21
 High quality sequence stop: 719.
 Location/Qualifiers
 1..938
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4518260"
 /clone.lib="NIH-MGC_91"
 /tissue_type="adenocarcinoma, cell_line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC library."
BASE COUNT 290 a 152 c 175 g 321 t
ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 938;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1867 TAGTCACTTCACAGCTCA 1886
 ||||||||||||||||||||
 Db 177 TAGTCACTTCACAGCTCA 158

RESULT 50
 AO688268
LOCUS AO688268
DEFINITION nbxb0077M07r CUGI Rice BAC Library Oryza sativa genomic clone
 nbxb0077M07r, DNA sequence.
ACCESSION AO688268
VERSION AO688268.1 GI:5329352
KEYWORDS GSS.
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 969)
 Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Rice Genome
 Unpublished (1998)
COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: GGAAACACTATGACCATG
 Class: BAC ends
 High quality sequence stop: 338.

FEATURES Location/Qualifiers
source 1..969
 /organism="Oryza sativa"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /clone="nbxb0077M07r"
 /clone.lib="CUGI Rice BAC library"
 /tissue_type="leaf"
 /lab_host="E. coli DH10B"
 /note="Vector: pBelosBAC11; Site_1: HindIII; Site_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocytiledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9%. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."
BASE COUNT 303 a 212 c 173 g 275 t 6 others
ORIGIN

Query Match 1.0%; Score 20; DB 12; Length 969;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 CAAGTAATTACCTGCTTAC 655
 ||||||||||||||||||||
 Db 218 CAAGTAATTACCTGCTTAC 237

Search completed: June 13, 2002, 11:34:17
 Job time: 10484 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 09:37:12 ; Search time 89.1 Seconds
(without alignments)
5510.902 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATCTACGACAAATAAAAA.....ATCTTCACACGACACACCA 1999

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size: 15

Total number of hits satisfying chosen parameters: 425

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1999	100.0	1999	3 US-08-961-083-1	Sequence 1, Appli
2	1438	71.9	2049	3 US-08-481-435-5	Sequence 5, Appli
3	796	39.8	960	2 US-08-245-511-3	Sequence 3, Appli
4	796	39.8	960	2 US-08-600-993A-3	Sequence 3, Appli
5	23	1.2	33	3 US-08-961-083-227	Sequence 227, App
6	19	1.0	2193	1 US-08-731-716-1	Sequence 1, Appli
7	19	1.0	2193	1 US-08-731-716-3	Sequence 3, Appli
8	18	0.9	4136	4 US-09-103-875-2	Sequence 2, Appli
9	17	0.9	27	3 US-08-961-083-228	Sequence 228, App
10	17	0.9	834	4 US-09-020-956-5	Sequence 5, Appli
11	17	0.9	834	4 US-09-030-607-5	Sequence 5, Appli
12	17	0.9	834	4 US-09-439-313-5	Sequence 5, Appli
13	17	0.9	971	4 US-08-630-915A-197	Sequence 197, App
14	17	0.9	1457	4 US-09-444-053-3	Sequence 3, Appli
15	16	0.8	20	4 US-09-444-053-5	Sequence 5, Appli
16	16	0.8	24	4 US-09-457-708-21	Sequence 21, Appli
17	16	0.8	30	1 US-08-232-144-2	Sequence 2, Appli
18	16	0.8	256	2 US-08-230-002-17	Sequence 17, Appli
19	16	0.8	256	2 US-08-678-854-17	Sequence 11, Appli
20	16	0.8	359	2 US-08-687-080-65	Sequence 65, Appli
21	16	0.8	574	1 US-08-592-126-137	Sequence 137, App
22	16	0.8	595	1 US-08-784-289-1	Sequence 1, Appli
23	16	0.8	643	4 US-08-858-207A-116	Sequence 116, App
24	16	0.8	766	4 US-08-858-207A-31	Sequence 31, Appli
25	16	0.8	875	4 US-08-936-165A-125	Sequence 125, App
26	16	0.8	890	1 US-08-592-126-78	Sequence 78, Appli
27	16	0.8	1186	1 US-08-368-236-2	Sequence 2, Appli

C 28	16	0.8	1295	1	US-08-245-295-6	Sequence 6, Appli
C 29	16	0.8	1295	1	US-08-481-130-6	Sequence 6, Appli
C 30	16	0.8	1295	1	US-08-656-984A-6	Sequence 6, Appli
C 31	16	0.8	1295	1	US-08-485-604-6	Sequence 6, Appli
C 32	16	0.8	1295	2	US-08-487-595-6	Sequence 6, Appli
C 33	16	0.8	1295	3	US-08-863-790-25	Sequence 25, Appli
C 34	16	0.8	1314	2	US-08-296-749-25	Sequence 25, Appli
C 35	16	0.8	1314	2	US-08-440-845D-9	Sequence 9, Appli
C 36	16	0.8	1353	3	US-08-868-458-9	Sequence 9, Appli
C 37	16	0.8	1353	2	US-08-611-280-24	Sequence 24, Appli
C 38	16	0.8	1353	4	US-09-195-940-44	Sequence 24, Appli
C 39	16	0.8	1400	2	US-08-001-078A-2	Sequence 2, Appli
C 40	16	0.8	1400	2	US-08-463-218-2	Sequence 2, Appli
C 41	16	0.8	1400	5	PCT-US94-00253-2	Sequence 2, Appli
C 42	16	0.8	1472	1	US-08-245-295-9	Sequence 9, Appli
C 43	16	0.8	1472	1	US-08-481-130-9	Sequence 9, Appli
C 44	16	0.8	1472	1	US-08-656-984A-9	Sequence 9, Appli
C 45	16	0.8	1472	1	US-08-485-604-9	Sequence 9, Appli
C 46	16	0.8	1472	2	US-08-487-595-9	Sequence 9, Appli
C 47	16	0.8	1500	1	US-08-443-568B-15	Sequence 15, Appli
C 48	16	0.8	1500	5	PCT-US94-06997-15	Sequence 15, Appli
C 49	16	0.8	1510	1	US-07-759-568-4	Sequence 4, Appli
C 50	16	0.8	1528	1	US-08-288-839-1	Sequence 1, Appli
C 51	16	0.8	1528	1	US-08-288-839-3	Sequence 3, Appli
C 52	16	0.8	1528	1	US-08-471-206-1	Sequence 1, Appli
C 53	16	0.8	1528	1	US-08-471-206-5	Sequence 5, Appli
C 54	16	0.8	1528	1	US-08-281-025D-1	Sequence 1, Appli
C 55	16	0.8	1547	1	US-08-288-839-4	Sequence 4, Appli
C 56	16	0.8	1547	1	US-08-471-206-8	Sequence 8, Appli
C 57	16	0.8	1568	4	US-08-884-077-3	Sequence 4, Appli
C 58	16	0.8	1605	6	5244676-1	Sequence 6, Appli
C 59	16	0.8	1639	2	US-08-468-819-77	Sequence 2, Appli
C 60	16	0.8	1727	6	5244676-4	Sequence 6, Appli
C 61	16	0.8	1748	1	US-08-202-056-8	Sequence 8, Appli
C 62	16	0.8	1755	1	US-07-912-900-29	Sequence 29, Appli
C 63	16	0.8	1755	1	US-08-285-309-29	Sequence 29, Appli
C 64	16	0.8	1755	2	US-08-502-046-29	Sequence 29, Appli
C 65	16	0.8	1801	1	US-08-557-917A-1	Sequence 1, Appli
C 66	16	0.8	1801	3	US-09-084-153-1	Sequence 1, Appli
C 67	16	0.8	1801	3	US-09-084-079-1	Sequence 1, Appli
C 68	16	0.8	1812	1	US-07-912-900-28	Sequence 28, Appli
C 69	16	0.8	1812	1	US-08-285-309-28	Sequence 28, Appli
C 70	16	0.8	1812	1	US-08-313-075A-37	Sequence 37, Appli
C 71	16	0.8	1812	2	US-08-502-046-28	Sequence 28, Appli
C 72	16	0.8	1824	3	US-08-606-505B-1	Sequence 1, Appli
C 73	16	0.8	1824	4	US-09-616-990-1	Sequence 1, Appli
C 74	16	0.8	1931	4	US-08-868-435-28	Sequence 28, Appli
C 75	16	0.8	1931	4	US-08-744-231-28	Sequence 28, Appli
C 76	16	0.8	2018	3	US-08-714-918-16	Sequence 16, Appli
C 77	16	0.8	2018	4	US-09-265-315-16	Sequence 16, Appli
C 78	16	0.8	2018	4	US-09-265-315-16	Sequence 16, Appli
C 79	16	0.8	2018	4	US-09-266-417-16	Sequence 16, Appli
C 80	16	0.8	2088	1	US-08-332-838-1	Sequence 1, Appli
C 81	16	0.8	2088	4	US-09-458-818B-3	Sequence 3, Appli
C 82	16	0.8	2214	1	US-08-245-295-7	Sequence 7, Appli
C 83	16	0.8	2214	1	US-08-481-130-7	Sequence 7, Appli
C 84	16	0.8	2214	1	US-08-656-984A-7	Sequence 7, Appli
C 85	16	0.8	2214	1	US-08-485-604-7	Sequence 7, Appli
C 86	16	0.8	2214	2	US-08-487-595-7	Sequence 7, Appli
C 87	16	0.8	2268	3	US-09-344-579-1	Sequence 1, Appli
C 88	16	0.8	2268	3	US-08-714-918-2	Sequence 2, Appli
C 89	16	0.8	2368	4	US-09-265-315-2	Sequence 2, Appli
C 90	16	0.8	2368	4	US-09-265-315-2	Sequence 2, Appli
C 91	16	0.8	2368	4	US-08-266-417-2	Sequence 2, Appli
C 92	16	0.8	2550	1	US-08-245-295-10	Sequence 10, Appli
C 93	16	0.8	2550	1	US-08-481-130-10	Sequence 10, Appli
C 94	16	0.8	2550	1	US-08-656-984A-10	Sequence 10, Appli
C 95	16	0.8	2550	1	US-08-485-604-10	Sequence 10, Appli
C 96	16	0.8	2550	2	US-08-487-595-10	Sequence 10, Appli
C 97	16	0.8	2936	4	US-08-943-731-64	Sequence 64, Appli
C 98	16	0.8	2988	1	US-08-245-295-1	Sequence 1, Appli
C 99	16	0.8	2988	1	US-08-481-130-1	Sequence 1, Appli
C 100	16	0.8	2988	1	US-08-656-984A-1	Sequence 1, Appli

C 101	16	0.8	2988	1	US-08-485-604-1	Sequence 1, Appl	C 174	15	0.8	610	4	US-09-328-111-209	Sequence 209, App
C 102	16	0.8	2988	1	US-08-487-595-1	Sequence 1, Appl	C 175	15	0.8	781	1	US-08-466-033-34	Sequence 34, Appl
C 103	16	0.8	3516	2	US-08-943-087-1	Sequence 1, Appl	C 176	15	0.8	781	1	US-08-444-733-34	Sequence 34, Appl
C 104	16	0.8	3635	1	US-08-252-626A-1	Sequence 1, Appl	C 177	15	0.8	781	2	US-08-464-134-34	Sequence 34, Appl
C 105	16	0.8	4144	1	US-08-218-686-1	Sequence 1, Appl	C 178	15	0.8	781	2	US-08-461-361-34	Sequence 34, Appl
C 106	16	0.8	4144	3	US-08-460-242-1	Sequence 1, Appl	C 179	15	0.8	781	2	US-08-485-910-34	Sequence 34, Appl
C 107	16	0.8	4248	3	US-08-678-614-1	Sequence 1, Appl	C 180	15	0.8	915	4	US-09-508-542-10	Sequence 10, Appl
C 108	16	0.8	4495	2	US-08-479-480A-3	Sequence 1, Appl	C 181	15	0.8	926	4	US-08-936-165A-118	Sequence 118, App
C 109	16	0.8	4565	3	US-08-776-265-1	Sequence 1, Appl	C 182	15	0.8	963	1	US-08-362-733-1	Sequence 1, Appl
C 110	16	0.8	4566	1	US-08-484-101B-37	Sequence 37, Appl	C 183	15	0.8	983	1	US-08-530-797-1	Sequence 1, Appl
C 111	16	0.8	4567	4	US-08-714-524D-37	Sequence 37, Appl	C 184	15	0.8	983	2	US-08-787-335-1	Sequence 1, Appl
C 112	16	0.8	4900	1	US-08-245-295-5	Sequence 5, Appl	C 185	15	0.8	1047	4	US-09-105-390-47	Sequence 47, Appl
C 113	16	0.8	4900	1	US-08-481-130-5	Sequence 5, Appl	C 186	15	0.8	1067	4	US-07-885-970A-2	Sequence 2, Appl
C 114	16	0.8	4900	1	US-08-656-984A-5	Sequence 5, Appl	C 187	15	0.8	1067	1	US-08-298-687A-2	Sequence 2, Appl
C 115	16	0.8	4900	1	US-08-485-604-5	Sequence 5, Appl	C 188	15	0.8	1067	1	US-08-241-943-2	Sequence 2, Appl
C 116	16	0.8	4900	2	US-08-487-595-5	Sequence 5, Appl	C 189	15	0.8	1067	1	US-08-378-588-2	Sequence 2, Appl
C 117	16	0.8	4900	2	US-08-863-790-26	Sequence 26, Appl	C 190	15	0.8	1067	1	US-08-298-829-2	Sequence 2, Appl
C 118	16	0.8	4900	3	US-08-296-749-26	Sequence 26, Appl	C 191	15	0.8	1067	2	US-08-811-094-2	Sequence 2, Appl
C 119	16	0.8	5077	1	US-08-245-295-8	Sequence 8, Appl	C 192	15	0.8	1067	5	PCT-US94-11121-2	Sequence 2, Appl
C 120	16	0.8	5077	1	US-08-481-130-8	Sequence 8, Appl	C 193	15	0.8	1131	4	US-09-105-390-63	Sequence 63, Appl
C 121	16	0.8	5077	1	US-08-656-984A-8	Sequence 8, Appl	C 194	15	0.8	1135	2	US-08-900-565-2	Sequence 2, Appl
C 122	16	0.8	5077	1	US-08-485-604-8	Sequence 8, Appl	C 195	15	0.8	1181	2	US-08-479-614-7	Sequence 7, Appl
C 123	16	0.8	5077	2	US-08-487-595-8	Sequence 8, Appl	C 196	15	0.8	1181	2	US-08-479-614-9	Sequence 9, Appl
C 124	16	0.8	5176	4	US-08-654-482-13	Sequence 13, Appl	C 197	15	0.8	1208	4	US-08-061-376-3	Sequence 3, Appl
C 125	16	0.8	5258	4	US-09-303-064-48	Sequence 48, Appl	C 198	15	0.8	1218	3	US-08-976-122-1	Sequence 1, Appl
C 126	16	0.8	5258	4	US-09-086-503-48	Sequence 48, Appl	C 199	15	0.8	1247	2	US-08-647-960-1	Sequence 1, Appl
C 127	16	0.8	5470	1	US-08-441-139-12	Sequence 12, Appl	C 200	15	0.8	1257	1	US-08-330-154-1	Sequence 1, Appl
C 128	16	0.8	5470	6	5196523-5	Patent No. 5196523	C 201	15	0.8	1260	2	US-08-839-581A-3	Sequence 3, Appl
C 129	16	0.8	12325	2	US-08-822-445-11	Sequence 11, Appl	C 202	15	0.8	1260	4	US-09-023-591A-3	Sequence 3, Appl
C 130	16	0.8	12225	2	US-09-396-540-11	Sequence 11, Appl	C 203	15	0.8	1274	3	US-08-335-844A-13	Sequence 13, Appl
C 131	16	0.8	12616	2	US-08-822-445-9	Sequence 9, Appl	C 204	15	0.8	1287	4	US-09-475-316A-67	Sequence 67, Appl
C 132	16	0.8	12616	2	US-09-396-540-9	Sequence 9, Appl	C 205	15	0.8	1290	5	US-08-012-543-1	Sequence 1, Appl
C 133	16	0.8	34303	2	US-08-735-609-4	Sequence 4, Appl	C 206	15	0.8	1290	5	PCT-US93-07645A-1	Sequence 1, Appl
C 134	16	0.8	34303	2	US-08-735-609-4	Sequence 4, Appl	C 207	15	0.8	1290	5	PCT-US93-07645A-1	Sequence 1, Appl
C 135	16	0.8	34303	3	US-09-315-372-4	Sequence 4, Appl	C 208	15	0.8	1290	5	US-08-371-121-15	Sequence 15, Appl
C 136	16	0.8	34303	3	US-09-244-752-4	Sequence 4, Appl	C 209	15	0.8	1313	2	US-08-463-911-6	Sequence 15, Appl
C 137	16	0.8	34303	3	US-09-245-497-4	Sequence 4, Appl	C 210	15	0.8	1365	4	US-08-858-207A-81	Sequence 81, Appl
C 138	16	0.8	34308	2	US-08-374-483-6	Sequence 6, Appl	C 211	15	0.8	1494	2	US-08-605-106-1	Sequence 2, Appl
C 139	16	0.8	35408	4	US-08-973-334-3	Sequence 3, Appl	C 212	15	0.8	1494	2	US-08-605-106-1	Sequence 1, Appl
C 140	16	0.8	35408	4	US-09-563-869A-3	Sequence 3, Appl	C 213	15	0.8	1488	1	US-08-638-911A-40	Sequence 40, Appl
C 141	16	0.8	35408	4	US-08-549-489-3	Sequence 3, Appl	C 214	15	0.8	1580	3	US-08-906-766-112	Sequence 112, App
C 142	16	0.8	35935	2	US-08-735-609-1	Sequence 1, Appl	C 215	15	0.8	1580	3	US-08-906-616-112	Sequence 112, App
C 143	16	0.8	35935	2	US-08-735-609-1	Sequence 1, Appl	C 216	15	0.8	1580	3	US-08-817-795-112	Sequence 112, App
C 144	16	0.8	35935	3	US-08-379-452-43	Sequence 43, Appl	C 217	15	0.8	1580	3	US-08-639-075A-112	Sequence 112, App
C 145	16	0.8	35935	3	US-09-315-372-1	Sequence 1, Appl	C 218	15	0.8	1580	4	US-09-012-431-112	Sequence 112, App
C 146	16	0.8	35935	3	US-09-244-752-1	Sequence 1, Appl	C 219	15	0.8	1580	4	US-09-012-692-112	Sequence 112, App
C 147	16	0.8	35935	3	US-09-245-497-1	Sequence 1, Appl	C 220	15	0.8	1580	4	US-08-906-613-112	Sequence 112, App
C 148	16	0.8	35935	3	US-09-409-670-43	Sequence 43, Appl	C 221	15	0.8	1580	5	PCT-US85-144A2A-112	Sequence 112, App
C 149	16	0.8	38682	4	US-08-943-731-2	Sequence 2, Appl	C 222	15	0.8	1600	2	US-08-933-750C-50	Sequence 50, Appl
C 150	16	0.8	72928	3	US-09-009-913-1	Sequence 1, Appl	C 223	15	0.8	1600	3	US-09-234-613-50	Sequence 50, Appl
C 151	16	0.8	176373	3	US-09-128-155-17	Sequence 17, Appl	C 224	15	0.8	1601	1	US-07-968-971A-10	Sequence 10, Appl
C 152	15	0.8	18	2	US-08-657-884-19	Sequence 19, Appl	C 225	15	0.8	1601	1	US-08-424-406-6	Sequence 6, Appl
C 153	15	0.8	18	2	US-09-158-980-19	Sequence 19, Appl	C 226	15	0.8	1618	1	US-07-885-970A-22	Sequence 22, Appl
C 154	15	0.8	19	1	US-08-513-846-8	Sequence 8, Appl	C 227	15	0.8	1618	1	US-08-298-687A-22	Sequence 22, Appl
C 155	15	0.8	25	4	US-08-638-931-22	Sequence 22, Appl	C 228	15	0.8	1618	1	US-08-298-829-22	Sequence 22, Appl
C 156	15	0.8	41	2	US-08-975-902-7	Sequence 7, Appl	C 229	15	0.8	1620	2	US-08-977-847-2	Sequence 2, Appl
C 157	15	0.8	41	3	US-09-251-565-7	Sequence 7, Appl	C 230	15	0.8	1620	2	US-09-195-021-2	Sequence 2, Appl
C 158	15	0.8	42	2	US-08-360-606B-24	Sequence 24, Appl	C 231	15	0.8	1647	1	US-07-885-970A-20	Sequence 20, Appl
C 159	15	0.8	99	1	US-07-968-971A-9	Sequence 9, Appl	C 232	15	0.8	1647	1	US-08-298-687A-20	Sequence 20, Appl
C 160	15	0.8	99	1	US-08-424-406-5	Sequence 5, Appl	C 233	15	0.8	1647	1	US-08-298-829-20	Sequence 20, Appl
C 161	15	0.8	121	4	US-08-817-787-2	Sequence 2, Appl	C 234	15	0.8	1668	1	US-07-885-970A-21	Sequence 21, Appl
C 162	15	0.8	121	4	US-08-817-787-9	Sequence 9, Appl	C 235	15	0.8	1668	1	US-08-298-687A-21	Sequence 21, Appl
C 163	15	0.8	128	2	US-08-687-080-102	Sequence 102, Appl	C 236	15	0.8	1668	1	US-08-298-829-21	Sequence 21, Appl
C 164	15	0.8	182	1	US-08-513-846-15	Sequence 15, Appl	C 237	15	0.8	1670	1	US-08-424-406-10	Sequence 10, Appl
C 165	15	0.8	184	1	US-08-513-846-16	Sequence 16, Appl	C 238	15	0.8	1672	1	US-08-530-797-13	Sequence 13, Appl
C 166	15	0.8	204	1	US-08-435-684A-60	Sequence 60, Appl	C 239	15	0.8	1672	2	US-08-787-339-13	Sequence 13, Appl
C 167	15	0.8	204	2	US-08-934-877A-60	Sequence 60, Appl	C 240	15	0.8	1716	3	US-09-029-267-1	Sequence 1, Appl
C 168	15	0.8	204	4	US-08-871-678C-60	Sequence 60, Appl	C 241	15	0.8	1721	3	US-08-857-212-2	Sequence 2, Appl
C 169	15	0.8	319	4	US-09-385-982-486	Sequence 486, Appl	C 242	15	0.8	1722	3	US-09-033-055A-1	Sequence 1, Appl
C 170	15	0.8	350	4	US-08-638-931-66	Sequence 66, Appl	C 243	15	0.8	1742	4	US-09-039-676-72	Sequence 2, Appl
C 171	15	0.8	425	4	US-09-342-681C-94	Sequence 94, Appl	C 244	15	0.8	1742	4	US-09-565-910-2	Sequence 2, Appl
C 172	15	0.8	492	4	US-09-030-607-197	Sequence 197, App	C 245	15	0.8	1744	1	US-08-424-406-7	Sequence 7, Appl
C 173	15	0.8	492	4	US-09-439-313-197	Sequence 197, App	C 246	15	0.8	1744	1	US-08-464-523B-7	Sequence 7, Appl

247	15	0.8	1750	1	US-08-415-823-1	Sequence 1, Appli	C 320	15	0.8	3561	2	US-08-590-554A-3	Sequence 3, Appli
248	15	0.8	1750	2	US-09-086-662-1	Sequence 1, Appli	C 321	15	0.8	3561	2	US-09-184-223-3	Sequence 3, Appli
249	15	0.8	1750	3	US-09-120-365-90	Sequence 90, Appli	C 322	15	0.8	3561	5	PCT-US93-12682-5	Sequence 5, Appli
250	15	0.8	1750	4	US-09-515-039-90	Sequence 90, Appli	C 323	15	0.8	3642	3	US-08-946-026-16	Sequence 16, Appli
251	15	0.8	1817	1	US-08-179-738-1	Sequence 1, Appli	C 324	15	0.8	3660	2	US-08-613-009A-5	Sequence 5, Appli
252	15	0.8	1817	2	US-08-628-145-1	Sequence 1, Appli	C 325	15	0.8	3670	2	US-08-898-038-1	Sequence 1, Appli
253	15	0.8	1827	3	US-08-824-878-2	Sequence 2, Appli	C 326	15	0.8	3707	4	US-09-276-531-42	Sequence 42, Appli
254	15	0.8	1827	2	US-09-353-688-2	Sequence 2, Appli	C 327	15	0.8	3713	1	US-08-100-709-1	Sequence 1, Appli
255	15	0.8	1833	1	US-08-179-738-6	Sequence 6, Appli	C 328	15	0.8	3713	1	US-08-176-865-1	Sequence 1, Appli
256	15	0.8	1833	2	US-08-628-145-6	Sequence 6, Appli	C 329	15	0.8	3713	2	US-08-474-038-1	Sequence 1, Appli
257	15	0.8	1833	3	US-09-515-039-6	Sequence 6, Appli	C 330	15	0.8	3713	1	US-08-779-046-1	Sequence 1, Appli
258	15	0.8	1833	4	US-09-515-039-6	Sequence 6, Appli	C 331	15	0.8	3713	2	US-08-881-340-1	Sequence 1, Appli
259	15	0.8	1856	2	US-08-360-606B-29	Sequence 29, Appli	C 332	15	0.8	3854	2	US-08-599-455B-42	Sequence 42, Appli
260	15	0.8	1862	1	US-08-179-738-4	Sequence 4, Appli	C 333	15	0.8	3854	4	US-09-066-781B-42	Sequence 42, Appli
261	15	0.8	1862	2	US-08-628-145-4	Sequence 4, Appli	C 334	15	0.8	3919	1	US-08-301-722A-1	Sequence 1, Appli
262	15	0.8	1863	3	US-09-305-381-1	Sequence 1, Appli	C 335	15	0.8	3979	4	US-09-085-199B-10	Sequence 10, Appli
263	15	0.8	1883	4	US-09-398-179-3	Sequence 3, Appli	C 336	15	0.8	3981	1	US-08-251-164-1	Sequence 1, Appli
264	15	0.8	1883	4	US-09-398-179-4	Sequence 4, Appli	C 337	15	0.8	3981	4	US-08-403-663-1	Sequence 1, Appli
265	15	0.8	1972	3	US-08-961-083-203	Sequence 203, App	C 338	15	0.8	4071	4	US-09-770-170-5	Sequence 5, Appli
266	15	0.8	2016	4	US-09-133-118-1	Sequence 1, Appli	C 339	15	0.8	4080	2	US-08-710-249-3	Sequence 3, Appli
267	15	0.8	2059	2	US-08-657-641-3	Sequence 3, Appli	C 340	15	0.8	4080	4	US-09-220-157A-3	Sequence 3, Appli
268	15	0.8	2059	5	PCT-US94-07233-3	Sequence 3, Appli	C 341	15	0.8	4098	4	US-09-268-866-1	Sequence 1, Appli
269	15	0.8	2137	1	US-08-444-005-16	Sequence 16, Appli	C 342	15	0.8	4201	1	US-08-080-235-4	Sequence 1, Appli
270	15	0.8	2158	1	US-07-602-608-1	Sequence 1, Appli	C 343	15	0.8	4201	3	US-08-465-713-4	Sequence 4, Appli
271	15	0.8	2158	1	US-08-261-578-1	Sequence 1, Appli	C 344	15	0.8	4201	5	PCT-US93-05857-4	Sequence 4, Appli
272	15	0.8	2167	3	US-08-884-324-7	Sequence 7, Appli	C 345	15	0.8	4517	4	US-09-140-804-9	Sequence 9, Appli
273	15	0.8	2184	4	US-09-105-390-8	Sequence 8, Appli	C 346	15	0.8	4643	2	US-08-605-106-6	Sequence 6, Appli
274	15	0.8	2187	2	US-08-679-635A-1	Sequence 1, Appli	C 347	15	0.8	5077	2	US-08-687-956A-22	Sequence 22, Appli
275	15	0.8	2242	3	US-09-400-742-1	Sequence 1, Appli	C 348	15	0.8	5361	4	US-08-973-462-2	Sequence 2, Appli
276	15	0.8	2242	3	US-08-618-651A-1	Sequence 1, Appli	C 349	15	0.8	5499	1	US-07-695-564-4	Sequence 4, Appli
277	15	0.8	2242	4	US-09-215-252-1	Sequence 1, Appli	C 350	15	0.8	5499	1	US-08-241-387-4	Sequence 4, Appli
278	15	0.8	2339	1	US-08-258-639A-1	Sequence 1, Appli	C 351	15	0.8	5629	1	US-07-695-564-2	Sequence 2, Appli
279	15	0.8	2339	2	US-08-900-951-1	Sequence 1, Appli	C 352	15	0.8	5629	1	US-08-241-387-2	Sequence 2, Appli
280	15	0.8	2339	4	US-08-858-207A-47	Sequence 47, Appli	C 353	15	0.8	5893	1	US-08-592-126-34	Sequence 34, Appli
281	15	0.8	2339	5	PCT-US95-07391A-1	Sequence 1, Appli	C 354	15	0.8	5893	2	US-08-687-080-44	Sequence 44, Appli
282	15	0.8	2388	1	US-08-638-911A-56	Sequence 56, Appli	C 355	15	0.8	5904	5	US-08-309-512-1	Sequence 1, Appli
283	15	0.8	2402	3	US-08-776-265-4	Sequence 4, Appli	C 356	15	0.8	5904	5	PCT-US92-087556A-1	Sequence 1, Appli
284	15	0.8	2446	2	US-08-061-376-1	Sequence 1, Appli	C 357	15	0.8	6152	4	US-08-973-462-1	Sequence 1, Appli
285	15	0.8	2446	2	US-09-016-000-9	Sequence 9, Appli	C 358	15	0.8	6371	3	US-08-836-325-13	Sequence 13, Appli
286	15	0.8	2617	3	US-09-161-443-1	Sequence 1, Appli	C 359	15	0.8	6404	3	US-08-836-325-14	Sequence 14, Appli
287	15	0.8	2809	3	US-08-448-194-5	Sequence 5, Appli	C 360	15	0.8	6830	2	US-08-822-445-1	Sequence 1, Appli
288	15	0.8	2809	4	US-08-867-921-5	Sequence 5, Appli	C 361	15	0.8	6830	4	US-09-396-540-1	Sequence 1, Appli
289	15	0.8	2899	2	US-08-624-581-2	Sequence 2, Appli	C 362	15	0.8	6850	4	US-09-298-367B-2	Sequence 2, Appli
290	15	0.8	2917	2	US-08-624-581-3	Sequence 3, Appli	C 363	15	0.8	7326	1	US-08-194-468-1	Sequence 1, Appli
291	15	0.8	2935	1	US-08-281-193-1	Sequence 1, Appli	C 364	15	0.8	7344	3	US-08-961-739-1	Sequence 1, Appli
292	15	0.8	2935	1	US-08-423-106-1	Sequence 1, Appli	C 365	15	0.8	7486	3	US-08-475-886-5	Sequence 5, Appli
293	15	0.8	2935	2	US-08-735-716-1	Sequence 1, Appli	C 366	15	0.8	7486	4	US-08-397-232-3	Sequence 3, Appli
294	15	0.8	2935	2	US-08-555-568B-1	Sequence 1, Appli	C 367	15	0.8	7488	3	US-08-475-886-3	Sequence 3, Appli
295	15	0.8	2935	4	US-09-519-223-1	Sequence 1, Appli	C 368	15	0.8	7493	3	US-08-475-886-1	Sequence 1, Appli
296	15	0.8	2935	5	PCT-US95-08069-1	Sequence 1, Appli	C 369	15	0.8	7493	4	US-08-397-232-1	Sequence 1, Appli
297	15	0.8	2959	2	US-08-624-581-1	Sequence 1, Appli	C 370	15	0.8	7493	4	US-09-127-387-1	Sequence 1, Appli
298	15	0.8	2971	1	US-07-718-575-7	Sequence 7, Appli	C 371	15	0.8	7833	1	US-08-074-879-9	Sequence 9, Appli
299	15	0.8	2971	1	US-08-481-206-7	Sequence 7, Appli	C 372	15	0.8	7833	1	US-08-468-057A-9	Sequence 9, Appli
300	15	0.8	2971	2	US-08-486-269A-7	Sequence 7, Appli	C 373	15	0.8	8532	1	US-08-452-655B-1	Sequence 1, Appli
301	15	0.8	3077	1	US-08-074-121-1	Sequence 1, Appli	C 374	15	0.8	8532	3	US-08-450-582-1	Sequence 1, Appli
302	15	0.8	3077	5	PCT-US94-06447-1	Sequence 1, Appli	C 375	15	0.8	8561	3	US-09-112-450-3	Sequence 3, Appli
303	15	0.8	3097	2	US-08-599-455B-1	Sequence 1, Appli	C 376	15	0.8	9046	5	PCT-US95-04682-1	Sequence 1, Appli
304	15	0.8	3097	4	US-09-069-781B-1	Sequence 1, Appli	C 377	15	0.8	9046	5	US-08-227-536-1	Sequence 1, Appli
305	15	0.8	3144	2	US-08-687-916-15	Sequence 15, Appli	C 378	15	0.8	9327	1	US-08-466-033-234	Sequence 234, App
306	15	0.8	3144	4	US-09-138-614-15	Sequence 15, Appli	C 379	15	0.8	9327	2	US-08-464-733-234	Sequence 234, App
307	15	0.8	3169	2	US-08-588-258B-27	Sequence 27, Appli	C 380	15	0.8	9327	2	US-08-464-733-14	Sequence 14, Appli
308	15	0.8	3169	3	US-08-460-505-27	Sequence 27, Appli	C 381	15	0.8	9327	2	US-08-464-733-14	Sequence 14, Appli
309	15	0.8	3169	5	PCT-US96-08295-27	Sequence 27, Appli	C 382	15	0.8	9327	2	US-08-485-910-234	Sequence 234, App
310	15	0.8	3188	4	US-08-943-731-183	Sequence 183, App	C 383	15	0.8	9391	1	US-08-638-911A-1	Sequence 1, Appli
311	15	0.8	3191	4	US-08-675-816-3	Sequence 3, Appli	C 384	15	0.8	9391	5	PCT-US95-06266-14	Sequence 14, Appli
312	15	0.8	3210	3	US-08-613-009A-6	Sequence 6, Appli	C 385	15	0.8	9392	1	US-08-466-033-14	Sequence 14, Appli
313	15	0.8	3222	3	US-08-613-009A-2	Sequence 2, Appli	C 386	15	0.8	9392	1	US-08-444-733-14	Sequence 14, Appli
314	15	0.8	3251	4	US-09-085-199B-6	Sequence 6, Appli	C 387	15	0.8	9392	2	US-08-464-733-14	Sequence 14, Appli
315	15	0.8	3286	2	US-08-363-124A-1	Sequence 1, Appli	C 388	15	0.8	9392	2	US-08-461-361-14	Sequence 14, Appli
316	15	0.8	3323	4	US-09-600-776-1	Sequence 1, Appli	C 389	15	0.8	9392	2	US-08-485-910-14	Sequence 14, Appli
317	15	0.8	3438	3	US-08-613-009A-1	Sequence 1, Appli	C 390	15	0.8	9606	1	US-07-741-940-1	Sequence 1, Appli
318	15	0.8	3561	1	US-08-485-568A-3	Sequence 3, Appli	C 391	15	0.8	9606	1	US-08-289-548A-1	Sequence 1, Appli
319	15	0.8	3561	1	US-08-357-698-5	Sequence 5, Appli	C 392	15	0.8	9606	1	US-08-452-654-1	Sequence 1, Appli


```
c 393 15 0.8 9606 2 US-08-370-235A-1 Sequence 1, Appl1
c 394 15 0.8 9687 4 US-09-133-944-2 Sequence 2, Appl1
c 395 15 0.8 9919 4 US-08-880-179-1 Sequence 1, Appl1
c 396 15 0.8 10322 4 US-09-330-330-3 Sequence 3, Appl1
c 397 15 0.8 10348 2 US-08-457-273B-41 Sequence 41, Appl1
c 398 15 0.8 10348 3 US-08-556-419-13 Sequence 13, Appl1
c 399 15 0.8 10348 4 US-09-041-886-14 Sequence 14, Appl1
c 400 15 0.8 10366 1 US-08-246-982A-5 Sequence 5, Appl1
c 401 15 0.8 10366 1 US-08-453-265-5 Sequence 5, Appl1
c 402 15 0.8 10803 3 US-09-080-044-1 Sequence 1, Appl1
c 403 15 0.8 11907 4 US-08-061-376-4 Sequence 4, Appl1
c 404 15 0.8 12847 1 US-08-550-715-1 Sequence 1, Appl1
c 405 15 0.8 14235 1 US-08-320-559-1 Sequence 1, Appl1
c 406 15 0.8 14235 1 US-08-327-392-1 Sequence 1, Appl1
c 407 15 0.8 14235 1 US-08-306-691B-55 Sequence 55, Appl1
c 408 15 0.8 14235 3 US-08-545-860D-1 Sequence 1, Appl1
c 409 15 0.8 14255 5 PCT-US94-04496-1 Sequence 1, Appl1
c 410 15 0.8 14507 3 US-08-785-150-1 Sequence 1, Appl1
c 411 15 0.8 14507 4 US-09-660-299-1 Sequence 1, Appl1
c 412 15 0.8 14507 4 US-09-435-377-1 Sequence 1, Appl1
c 413 15 0.8 19227 3 US-09-090-793-13 Sequence 13, Appl1
c 414 15 0.8 20084 4 US-08-943-731-5 Sequence 5, Appl1
c 415 15 0.8 22846 2 US-08-469-461-3 Sequence 3, Appl1
c 416 15 0.8 22846 3 US-07-890-609-3 Sequence 3, Appl1
c 417 15 0.8 28994 3 US-08-884-324-14 Sequence 14, Appl1
c 418 15 0.8 35100 1 US-08-306-691B-19 Sequence 19, Appl1
c 419 15 0.8 35100 5 PCT-US93-06251-19 Sequence 19, Appl1
c 420 15 0.8 40138 3 US-09-090-793-12 Sequence 12, Appl1
c 421 15 0.8 48974 4 US-08-920-422-17 Sequence 17, Appl1
c 422 15 0.8 51259 3 US-08-781-891-209 Sequence 209, App
c 423 15 0.8 51952 3 US-08-947-823-1 Sequence 1, Appl1
c 424 15 0.8 90050 4 US-09-245-041-5 Sequence 5, Appl1
c 425 15 0.8 441529 4 US-09-103-840A-1 Sequence 1, Appl1
```

ALIGNMENTS

```
RESULT 1
US-08-961-083-1
: Sequence 1, Application US/08961083
: Patent No. 6159469
: GENERAL INFORMATION:
: APPLICANT: Choi et. al.
: TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
: NUMBER OF SEQUENCES: 452
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,083
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 1:
```

```
SEQUENCE CHARACTERISTICS:
: LENGTH: 1999 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
US-08-961-083-1

Query Match          100.0%; Score 1999; DB 3; Length 1999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAATCTAGACAAATAAATCACTAGTGTGACTGGTTCGTAAGCGCGGTCAA 60
DB 1 TAAATCTAGACAAATAAATCACTAGTGTGACTGGTTCGTAAGCGCGGTCAA 60
QY 61 TGCCCAAGCTAATGATATTCACAGATTTGGTTAAGGCATGTTTCTATCGAAGCCA 120
DB 61 TGCCCAAGCTAATGATATTCACAGATTTGGTTAAGGCATGTTTCTATCGAAGCCA 120
QY 121 TCGCTTCTTGACACACAGGGGGAATGATACATCCGATCTTGAGACTTTTGGCGAA 180
DB 121 TCGCTTCTTGACACACAGGGGGAATGATACATCCGATCTTGAGACTTTTGGCGAA 180
QY 181 TCTGCAAAAGCAATTCCTCCAGGTGATCACTCTCACCACAGTTGATTAAGTTGAC 240
DB 181 TCTGCAAAAGCAATTCCTCCAGGTGATCACTCTCACCACAGTTGATTAAGTTGAC 240
QY 241 TTACTTTTCAACTTCGACCTCGACAGACTATTTTCGTAAGGCTCAGGAACCTGGTT 300
DB 241 TTACTTTTCAACTTCGACCTCGACAGACTATTTTCGTAAGGCTCAGGAACCTGGTT 300
QY 301 AGGATTTCACTGTGAACAAACCAACCAAGCAAGAAATCTGACCTATATAATA 360
DB 301 AGGATTTCACTGTGAACAAACCAACCAAGCAAGAAATCTGACCTATATAATA 360
QY 361 GGTCTACATGCTTAATGGAGCACTATGATGACAGACAGCTCAAAACTATATGGTAA 420
DB 361 GGTCTACATGCTTAATGGAGCACTATGATGACAGACAGCTCAAAACTATATGGTAA 420
QY 421 AGACCTCAATATTAATTAAGTTTACCTCAGTTAGCTTGGCTGGGAATGGCTAGGCACC 480
DB 421 AGACCTCAATATTAATTAAGTTTACCTCAGTTAGCTTGGCTGGGAATGGCTAGGCACC 480
QY 481 AAACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGAGCGGAACTTGGTCT 540
DB 481 AAACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGAGCGGAACTTGGTCT 540
QY 541 ATCTGAATGAATAAATCAAGGCTACATCTGCTGAACAGTATGAGAAAGCACTAATAC 600
DB 541 ATCTGAATGAATAAATCAAGGCTACATCTGCTGAACAGTATGAGAAAGCACTAATAC 600
QY 541 ATCTGAATGAATAAATCAAGGCTACATCTGCTGAACAGTATGAGAAAGCACTAATAC 600
DB 541 ATCTGAATGAATAAATCAAGGCTACATCTGCTGAACAGTATGAGAAAGCACTAATAC 600
QY 601 ACCAATTACTGATGAGATCAAAAGTCAATAGCAAGTAATTAACCTGCTACATGGA 660
DB 601 ACCAATTACTGATGAGATCAAAAGTCAATAGCAAGTAATTAACCTGCTACATGGA 660
QY 661 TAAATACCTCAAGAGTCAATCAATCAAGTTGAAGAAGAAACAGGCTAATACCTACTCAC 720
DB 661 TAAATACCTCAAGAGTCAATCAATCAAGTTGAAGAAGAAACAGGCTAATACCTACTCAC 720
QY 721 AACTGGAGTGAATGCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
DB 721 AACTGGAGTGAATGCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
QY 781 TTACATATACAGAGATTAATGCTTGCATCCAGAGATGATTCGCAAGTGCCTTCCAT 840
DB 781 TTACATATACAGAGATTAATGCTTGCATCCAGAGATGATTCGCAAGTGCCTTCCAT 840
QY 841 TGTGATGTTTCTAAGGTTAAAGTCAATGTCGCGAGCTAGAGACAGGCATCAGTCAAGTAA 900
DB 841 TGTGATGTTTCTAAGGTTAAAGTCAATGTCGCGAGCTAGAGACAGGCATCAGTCAAGTAA 900
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCGGACTGGGATCAACTATGAA 960
```



```

Db 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAAAAACAACCGCGATCGGATCAACTATGAA 960
QY 961 ACCGATTCACAGACTATGCTCCCGCTTGAGTAGCGTCTACAGTTCAACTGCTACTAT 1020
Db 961 ACCGATTCACAGACTATGCTCCCGCTTGAGTAGCGTCTACAGTTCAACTGCTACTAT 1020
QY 1021 CGTTCACGATGAGCCCTTAATACCTCGGACAATACCTGCTGTTTAACTGGGATAG 1080
Db 1021 CGTTCACGATGAGCCCTTAATACCTCGGACAATACCTGCTGTTTAACTGGGATAG 1080
QY 1081 GGGCTACTTGGCAACATACCTTGCATATAGCCCTGCAACAATCGGAAAAGTCCAGC 1140
Db 1081 GGGCTACTTGGCAACATACCTTGCATATAGCCCTGCAACAATCGGAAAAGTCCAGC 1140
QY 1141 CGTGGAACTCTAAACAGAGTGGGACATACCGGCGCAAGCTTCCCAATGGGCTAGG 1200
Db 1141 CGTGGAACTCTAAACAGAGTGGGACATACCGGCGCAAGCTTCCCAATGGGCTAGG 1200
QY 1201 AATGCACTACCCCAAGTATTCCTACTCAATGCCATTTCAAGTAACACACCGAATCAGA 1260
Db 1201 AATGCACTACCCCAAGTATTCCTACTCAATGCCATTTCAAGTAACACACCGAATCAGA 1260
QY 1261 CAAAAAATATGAGCAAGCACTAGTGAAGATGGCTGCTTACGCTGCTTGCAGATGG 1320
Db 1261 CAAAAAATATGAGCAAGCACTAGTGAAGATGGCTGCTTACGCTGCTTGCAGATGG 1320
QY 1321 TGGAACTTACTTAATACCAATGATATATCCATTAAGTCTCTTATGATGGAGTGAAGA 1380
Db 1321 TGGAACTTACTTAATACCAATGATATATCCATTAAGTCTCTTATGATGGAGTGAAGA 1380
QY 1381 AGAGTTCTCTAATGTGCGAAGTCTGCGCATGAGGAAAGCAAGCGCTTATGATGACCGA 1440
Db 1381 AGAGTTCTCTAATGTGCGAAGTCTGCGCATGAGGAAAGCAAGCGCTTATGATGACCGA 1440
QY 1441 CATGATGAAAACAGTCTTGACTTATGAGTGAAGTGAAGAAATGCTTCTTGGCTCCC 1500
Db 1441 CATGATGAAAACAGTCTTGACTTATGAGTGAAGTGAAGAAATGCTTCTTGGCTCCC 1500
QY 1501 TCAGGCTGGTAAACAGAGACCTCTAATATACAGAGAGGAAATGGAATCCACATGAA 1560
Db 1501 TCAGGCTGGTAAACAGAGACCTCTAATATACAGAGAGGAAATGGAATCCACATGAA 1560
QY 1561 GACCTCTCAATTTTGTAGACCTGATGAATCTTTGCTGCTATAGCGCTAAATATTCAT 1620
Db 1561 GACCTCTCAATTTTGTAGACCTGATGAATCTTTGCTGCTATAGCGCTAAATATTCAT 1620
QY 1621 GGGCTGTATGACAGGCTATTTCTAACCCTGTGACACACACTTGTAGGCAATGGCCTTACG 1680
Db 1621 GGGCTGTATGACAGGCTATTTCTAACCCTGTGACACACACTTGTAGGCAATGGCCTTACG 1680
QY 1681 CGCTGCCAAATTTTACCGCTCTATGATGATGATGATGATGATGATGATGATGATGATG 1740
Db 1681 CGCTGCCAAATTTTACCGCTCTATGATGATGATGATGATGATGATGATGATGATGATG 1740
QY 1741 TTGGAATATACAGAGGCGCTCTACAGAAATGGAATTCGATTTAAATTTGGGCTCG 1800
Db 1741 TTGGAATATACAGAGGCGCTCTACAGAAATGGAATTCGATTTAAATTTGGGCTCG 1800
QY 1801 TTCTACGTGAGTACCTGCTCTCAACAACACCCCATCACTGAAGTTCAGGCTCATC 1860
Db 1801 TTCTACGTGAGTACCTGCTCTCAACAACACCCCATCACTGAAGTTCAGGCTCATC 1860
QY 1861 ATCAGATATGTTCAACTTACAGTCTAGCTCAACACTCTCAAGCACAATTAATAGTACGAC 1920
Db 1861 ATCAGATATGTTCAACTTACAGTCTAGCTCAACACTCTCAAGCACAATTAATAGTACGAC 1920
QY 1921 TACCAATCTTAACATATATAGCAACAATTCATAATACACCCCTGATCAACAAATTCAGAA 1980
Db 1921 TACCAATCTTAACATATATAGCAACAATTCATAATACACCCCTGATCAACAAATTCAGAA 1980
QY 1981 TCCTCAACAGCACAACCA 1999

```

```

Db 1981 TCCTCAACAGCACAACCA 1999

RESULT 2
US-08-481-435-5
; Sequence 5, Application US/08481435
; Patent No. 6027906
; GENERAL INFORMATION:
; APPLICANT: Balganes, Tanjore S
; APPLICANT: Town, Christine
; TITLE OF INVENTION: No. 6027906el Polypeptides
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESS: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481.435
; FILING DATE: 10-JUL-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: IN 580/MAS/94
; FILING DATE: 01-JUL-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: SE 9404072-2
; FILING DATE: 24-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stener, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: pm 1
; IMMEDIATE SOURCE:
; LIBRARY: PCR cloning
; CLONE: PARC 0512 Soluble PBP 1A del 38
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2049
; NAME/KEY: mat_peptide
; LOCATION: 1..2046
; US-08-481-435-5

```

```

Query Match 71.9%; Score 1438; DB 3; Length 2049;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTACAGCAATATAATCAATCATTCGCTGATGGCTCTGAAACCGCGGTCAA 60
Db 48 TAAATCTACAGCAATATAATCAATCATTCGCTGATGGCTCTGAAACCGCGGTCAA 107
QY 61 TGCCCAAGCTAATATATTTCCACAGATTTGGTTAAGCAATCGTTCTATGAAAGACA 120

```


Db	108	TCGCCAGCTAATATATATATTCGCCACAGATTGGTTAAAGCAATCGTTTCTATCGAAGACCA	167
QY	121	TCGCTCTTCGACACAGAGGGGATGTGATACCATCGGTATCTGGAGCTTCTTCGCGAA	180
Db	168	TCGCTCTTCGACACAGAGGGGATGTGATACCATCGGTATCTGGAGCTTCTTCGCGAA	227
QY	181	TCGCAAGCAATTCCTCCAAAGTGGATCACTCCACCACAGTTATTAATTGAC	240
Db	228	TCGCAAGCAATTCCTCCAAAGTGGATCACTCCACCACAGTTATTAATTGAC	287
QY	241	TTACCTTTCAACTTCGACTTCGACCTCCGACAGCATATTTCTCGTAAGGCTCAGAGCTTGGTT	300
Db	288	TTACTTTTCAACTTCGACTTCGACCTCCGACAGCATATTTCTCGTAAGGCTCAGAGAGCTTGGTT	347
QY	301	AGCGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATPA	360
Db	348	AGCGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATPA	407
QY	361	GGTCTACATGTCTAATGGGAATGTGAAATGCAAGACAGCTCAAAACTACTATATGTNA	420
Db	408	GGTCTACATGTCTAATGGGAATGTGAAATGCAAGACAGCTCAAAACTACTATATGTNA	467
QY	421	AGACCTCAATTAATTAAGTTTACCTCGTAGGCTTGCTGGCTGGAATCCTCAGAGCAC	480
Db	468	AGACCTCAATTAATTAAGTTTACCTCGTAGGCTTGCTGGCTGGAATCCTCAGAGCAC	527
QY	481	AAACCAATATGACCCCTATTTCACATCCAGAGGACCCCAAGACCGCGAAACTTGGTCTT	540
Db	528	AAACCAATATGACCCCTATTTCACATCCAGAGGACCCCAAGACCGCGAAACTTGGTCTT	587
QY	541	ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAAACGATATGAAAGCAAGTAAATAC	600
Db	588	ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAAACGATATGAAAGCAAGTAAATAC	647
QY	601	ACCAATTACTATGAGACTACAGAAAGTCCAAATCAGACAAATTAATACCGCTTACATGGA	660
Db	648	ACCAATTACTATGAGACTACAGAAAGTCCAAATCAGACAAATTAATACCGCTTACATGGA	707
QY	661	TAAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAAGCAAGGCTATTAACCTACTAC	720
Db	708	TAAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAAGCAAGGCTATTAACCTACTAC	767
QY	721	AACGGGATGGATGTCTACACAAATGTAGACCAAGAAAGCTCAAAAACATCTGGGAAAT	780
Db	768	AACGGGATGGATGTCTACACAAATGTAGACCAAGAAAGCTCAAAAACATCTGGGAAAT	827
QY	781	TTACAAATCACAAGCAATACGTTGGCTTCCAGAGCATGGAATTGCAAGTGGCTTACCAT	840
Db	828	TTACAAATCACAAGCAATACGTTGGCTTCCAGAGCATGGAATTGCAAGTGGCTTACCAT	887
QY	841	TGTTGATGTTTCTAAGCGTAAAGTCATTTGGCCAGCTAGGAGCACGCCATCACTCAAGTAA	900
Db	888	TGTTGATGTTTCTAAGCGTAAAGTCATTTGGCCAGCTAGGAGCACGCCATCACTCAAGTAA	947
QY	901	TGTTTCCCTGGGAATTTAACCAAGCAGTAAAGAAACAAACCGCGACTGGGGATCAACTATGAA	960
Db	948	TGTTTCCCTGGGAATTTAACCAAGCAGTAAAGAAACAAACCGCGACTGGGGATCAACTATGAA	1007
QY	961	ACCGATCACAAGATATCTCTCGCTGGAGTACGGGTGTCTACGANTCAACGCTACTAT	1020
Db	1008	ACCGATCACAAGATATCTCTCGCTGGAGTACGGGTGTCTACGAGTCAACTGCCACTAT	1067
QY	1021	CGTTACAGATAGCCCTTAATACTACCTGGGACAAATACTCTGTTTAACTGGGATAG	1080
Db	1068	CGTTACAGATAGCCCTTAATACTACCTGGGACAAATACTCTGTTTAACTGGGATAG	1127
QY	1081	GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAATGCGAAGCTCCAGC	1140
Db	1128	GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAATGCGAAGCTCCAGC	1187
QY	1141	CGTGGAAACCTTAACAAGGTGGAGCTCAACCGCGCAAGACTTTCCTAAATGCTCTAG	1200
Db	1188	CGTGGAAACCTTAACAAGGTGGAGCTCAACCGCGCAAGACTTTCCTAAATGCTCTAG	1247

QY	1201	AAAGCAGTACCCAGTATTTACATACATCAATGGCAATTTTCAAGTACACACACCCAGTACGA	1260
Db	1248	AATGACTACCAAGTATATCACTCATAAATGSCATTTTCAAGTACACACACCGAATTCGA	1307
QY	1261	CAAAAAATATGACCAAGTAGTGAAGAGATGGCTGCTGCTTACGCTGCTTGGCAAAATG	1320
Db	1308	CAAAAAATATGAGACCAAGTAGTGAAGAAAGATGGCTGCTTACGCTGCTTGGCAAAATG	1367
QY	1321	TGGAACTTACTATTAACCAATGTATATCCATAAAGTCGCTTTAGTGAATGGAGTGA	1380
Db	1368	TGGAACTTACTATTAACCAATGTATATCCATAAAGTCGCTTTAGTGAATGGAGTGA	1427
QY	1381	AGAGTTCCTCTAATGTCGAACTGCTGCGATGAAGAAAGACAGACCTATATGATACCGA	1440
Db	1428	AGAGTTCCTCTAATGTCGAACTGCTGCGATGAAGAAAGACAGACCTATATGATACCGA	1487
QY	1441	CATGATGAAAAACAGTCTTGAATTATGGAACGTGAGCAAAATGCCATCTTCTTGGCTCC	1500
Db	1488	CATGATGAAAAACAGTCTTGAATTATGGAACGTGAGCAAAATGCCATCTTCTTGGCTCC	1547
QY	1501	TCAGGCTGTGTAAGACAGAACCTCTAACTTACAGACGAGGAATTTGAAACACACATCA	1566
Db	1548	TCAGGCTGTGTAAGACAGAACCTCTAACTTACAGACGAGGAATTTGAAACACACATCA	1607
QY	1561	GACCTCTCAATTTTGTACACCTGATGAACTAATTTGCTGCTATACGCGTAATATTCAAT	1620
Db	1608	GACCTCTCAATTTTGTACACCTGATGAACTAATTTGCTGCTATACGCGTAATATTCAAT	1667
QY	1621	GCGTGTATGACAGGCTATTTCTAACCCTGACACACCTTGTAGGCAATGGCTTAGCGT	1680
Db	1668	GCGTGTATGACAGGCTATTTCTAACCCTGACACACCTTGTAGGCAATGGCTTAGCGT	1727
QY	1681	CGCTGCCAAATTTACGCGCTCTATGATGACCTACCTGCTGGAAGGAAGCAATCCGAA	1740
Db	1728	CGCTGCCAAATTTACGCGCTCTATGATGACCTACCTGCTGGAAGGAAGCAATCCGAA	1787
QY	1741	TTGGAATATACCAAGGAGGCTCTACAGAAATGGAGAAATCGTATTTAAAAATGGTCTCG	1800
Db	1788	TTGGAATATACCAAGGAGGCTCTACAGAAATGGAGAAATCGTATTTAAAAATGGTCTCG	1847
QY	1801	TTTCTACGTGGAACCTACCTGCTGCACACACACCCCATTAACCTGAAGATTCAGCTATC	1860
Db	1848	TTTCTACGTGGAACCTACCTGCTGCACACACACCCCATTAACCTGAAGATTCAGCTATC	1907
QY	1861	ATCGATATGTCACCTTACAGAGTCTAGCTCAACCATCCACACACAAATTAATAGTAGAC	1920
Db	1908	ATCGATATGTCACCTTACAGAGTCTAGCTCAACCATCCACACACAAATTAATAGTAGAC	1967
QY	1921	TACCACTCTAACATTAATAGGACAAATCAATATACMACCCCTGATCAACAAATCAGA	1980
Db	1968	TACCACTCTAACATTAATAGGACAAATCAATATACMACCCCTGATCAACAAATCAGA	2027
QY	1981	TCCTCAGACGACACACCA 1999	
Db	2028	TCCTCAGACGACACACCA 2046	
RESULT 3			
US-08-245-511-3			
; Sequence 3, Application US/08245511			
; Patent No. 5928900			
; GENERAL INFORMATION:			
; APPLICANT: Measure, H Robert			
; APPLICANT: Pearce, Barbara J			
; APPLICANT: Tuomanen, Elaine			
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND			
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THERON			
; NUMBER OF SEQUENCES: 58			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Klauber & Jackson			
; STREET: 411 Hackensack Avenue			
; CITY: Hackensack			

RESULT 3
US-08-245-511-3
Sequence 3, Application US/08245511
Patent No. 5928900
GENERAL INFORMATION:
APPLICANT: Masure, H Robert
APPLICANT: Pearce, Barbara J
APPLICANT: Tuomari, Elaine
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack


```
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/245,511
FILING DATE: 18-MAY-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/116,541
FILING DATE: 01-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-069 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 960 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
STRAIN: R6
IMMEDIATE SOURCE:
CLONE: SPR042
FEATURE:
NAME/KEY: CDS
LOCATION: 1..960
US-08-245-511-3

Query Match          39.8%; Score 796; DB 2; Length 960;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TAAATCTACGACAAATTAATCAATCATCTGCTGCTGGTCTGACCGCGCTCAA 60
DB 12 TAAATCTACGACAAATTAATCAATCATCTGCTGCTGGTCTGACCGCGCTCAA 71
OY 61 TGGCCAACTATGATTTCCACAGATTGGTTAAAGCAATCCTTTCTATCGAAGACCA 120
DB 72 TGGCCAACTATGATTTCCACAGATTGGTTAAAGCAATCCTTTCTATCGAAGACCA 131
OY 121 TCGCTTCTCGACACAGAGGGGATGATAGCATCCGATCCGTGGAGGCTTTCTTGGCAA 180
DB 132 TCGCTTCTCGACACAGAGGGGATGATAGCATCCGATCCGTGGAGGCTTTCTTGGCAA 191
OY 181 TGTGAAAGCAATTCCTTCACAGGTGATCAACTCAACCAAGTTGATTAGTTGAC 240
DB 192 TGTGAAAGCAATTCCTTCACAGGTGATCAACTCAACCAAGTTGATTAGTTGAC 251
OY 241 TTACTTTCACTTCGACCTCCGACACAGATTTCTGTAAGGCTCAGGAAGCTTGCTT 300
DB 252 TTACTTTCACTTCGACCTCCGACACAGATTTCTGTAAGGCTCAGGAAGCTTGCTT 311
OY 301 AGCGATTAGTTAGAAACAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATA 360
DB 312 AGCGATTAGTTAGAAACAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATA 371
OY 361 GGTCTACATGCTATATGGAACTATGGAATGACAGACAGCTCAAACTACTATGTGTA 420
DB 420 GGTCTACATGCTATATGGAACTATGGAATGACAGACAGCTCAAACTACTATGTGTA 431
```

```
DB 372 GGTCTACATGCTATATGGAACTATGGAATGACAGACAGCTCAAACTACTATGTGTA 431
OY 421 AGACCTCAATATATTAAGTTTACCTCAGTTAGCTTGGTGGTGAATGCCACAGACACC 480
DB 432 AGACCTCAATATATTAAGTTTACCTCAGTTAGCTTGGTGGTGAATGCCACAGACACC 491
OY 481 AAACCAATATGACCCCTATTTACATCCAGACAGCCCAAGACCCGCAAACTTGCTTT 540
DB 492 AAACCAATATGACCCCTATTTACATCCAGACAGCCCAAGACCCGCAAACTTGCTTT 551
OY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCACTAATAC 600
DB 552 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCACTAATAC 611
OY 601 ACCAATTCGATGAGCTACAAAGCTCAAAAGCTCAAAAGCTAATTCCTGCTTACATGGA 660
DB 612 ACCAATTCGATGAGCTACAAAGCTCAAAAGCTCAAAAGCTAATTCCTGCTTACATGGA 671
OY 661 TAATTACCTCAAGAAAGTCAATCAAGTTGAAGAAAGCAAGGCTATTAACCTACTCAC 720
DB 672 TAATTACCTCAAGAAAGTCAATCAAGTTGAAGAAAGCAAGGCTATTAACCTACTCAC 731
OY 721 AACTGGGATGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTTGGAATAT 780
DB 732 AACTGGGATGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTTGGAATAT 791
OY 781 TTACAAATCAGACGAATACGCTGCTATCCAGACATGAATTCGAATGCTTACCAT 840
DB 792 TTACAAATCAGACGAATACGCTGCTATCCAGACATGAATTCGAATGCTTACCAT 851
OY 841 TGTGATGCTTTTAAAGCGTAAGTATCCAGTAGAGCAGCCATCACTCAAGTAA 900
DB 852 TGTGATGCTTTTAAAGCGTAAGTATCCAGTAGAGCAGCCATCACTCAAGTAA 911
OY 901 TGTTCCTTGGGAATTAACCAAGCAGTAGAACAACCGCGACTGGGGA 949
DB 912 TGTTCCTTGGGAATTAACCAAGCAGTAGAACAACCGCGACTGGGGA 960

RESULT 4
US-08-600-993A-3
; Sequence 3, Application US/08600993A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Measure, H Robert
; APPLICANT: Pearce, Barbara J
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACETILULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
```


ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 600-1-069 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201 487-5800
 TELEFAX: 201 343-1684
 TELEX: 133521
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 960 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Streptococcus pneumoniae
 STRAIN: R6
 IMMEDIATE SOURCE:
 CLONE: SPR042
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..960
 US-08-961-083-1.0115-3

Query Match 39.8%; Score 796; DB 2; Length 960;
 Best Local Similarity 99.7%; Pred. No. 0;

Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 TAAATCTACGACATAAAATCACTATTGCTGACCTTGGTTGCAAGCGCGCA 60
 12 TAAATCTACGACATAAAATCACTATTGCTGACCTTGGTTGCAAGCGCGCA 71
 61 TGGCCAGCTAATGATATTCCTCCACAGATTTGTTAAGCAATCGTTCTATGCAAGCA 120
 72 TGGCCAGCTAATGATATTCCTCCACAGATTTGTTAAGCAATCGTTCTATGCAAGCA 131
 121 TGGCTTCTGACACAGAGGGGATGATACCATCCGATCCCTGGAGCTTCTTGCGCA 180
 132 TGGCTTCTGACACAGAGGGGATGATACCATCCGATCCCTGGAGCTTCTTGCGCA 191
 181 TGTGCAAGCAATTCCTCCACAGATTTGTTAAGCAATCGTTCTATGCAAGCA 240
 192 TGTGCAAGCAATTCCTCCACAGATTTGTTAAGCAATCGTTCTATGCAAGCA 251
 241 TTTACTTTTCACTTGCACCTTCGACACAGACTATTTCTGTAAGGCTCAGGAAGCTTG 300
 252 TTTACTTTTCACTTGCACCTTCGACACAGACTATTTCTGTAAGGCTCAGGAAGCTTG 311
 301 AGGATTCAGTTGACCAAAAGCAACCAAGCAAGAAATCTGACCTACTATTAATA 360
 312 AGGATTCAGTTGACCAAAAGCAACCAAGCAAGAAATCTGACCTACTATTAATA 371
 361 GGTCTACATGTCTAATGGGAGTATGGAATGCAGACAGCTCAAAAACACTATGTTAA 420
 372 GGTCTACATGTCTAATGGGAGTATGGAATGCAGACAGCTCAAAAACACTATGTTAA 431
 421 AGACCTCAATTAATTTAAGTTACCTCAGTTAGCTTGGTGGTGAATGGCTCAGGAC 480
 432 AGACCTCAATTAATTTAAGTTACCTCAGTTAGCTTGGTGGTGAATGGCTCAGGAC 491
 481 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGAGCCGCAAACTTGCTT 540
 492 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGAGCCGCAAACTTGCTT 551
 541 ATCTGAAATGAAATCAAGGCTACATCTGCTGTAACATGAAAGCAATCAATCAATAC 600
 552 ATCTGAAATGAAATCAAGGCTACATCTGCTGTAACATGAAAGCAATCAATCAATAC 611
 601 ACCAATTACTGATGACTACAAAGTCTCAAAATCAGCAAGTAATACCTTGCTTACATGA 660

|||||
 Db 612 ACCAATTACTGATGGGCTACCAAGTCTCAATCAGCAAGTAATACCTTGCTTACATGA 671
 661 TAATTTACTTAAGGAATGATCAATCAAGTTGAAGGAAGCAAGGCTATTAACCTACTCAC 720
 672 TAATTTACTTAAGGAATGATCAATCAAGTTGAAGGAAGCAAGGCTATTAACCTACTCAC 731
 721 AACTGGATGATGCTTACACCAATGATAGACCAAGGAAGTCAAAAACATCTGTTGATAT 780
 732 AACTGGATGATGCTTACACCAATGATAGACCAAGGAAGTCAAAAACATCTGTTGATAT 791
 781 TTACAATACAGACGAATAGCTGCTTATCCAGACGATGAATTCGAAGCTTCTTACCAT 840
 792 TTACAATACAGACGAATAGCTGCTTATCCAGACGATGAATTCGAAGCTTCTTACCAT 851
 841 TGTGATGTTTCTACGAGTAAGTCAATGCTGCGCAGCTAGAGACCGCATGCTCAAGTAA 900
 852 TGTGATGTTTCTACGAGTAAGTCAATGCTGCGCAGCTAGAGACCGCATGCTCAAGTAA 911
 901 TGTTCCTTGGGAATTAACCAAGCAGTAGAAGCAAAACCGGACTGGGGA 949
 912 TGTTCCTTGGGAATTAACCAAGCAGTAGAAGCAAAACCGGACTGGGGA 960

RESULT 5
 US-08-961-083-227
 ; Sequence 227, Application US/08961083
 ; Patent No. 6159469

GENERAL INFORMATION:
 APPLICANT: Choi et. al.
 TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
 NUMBER OF SEQUENCES: 452
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage

COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/961,083
 FILING DATE:

CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:

ATTORNEY/AGENT INFORMATION:
 NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB340P2

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 227:

SEQUENCE CHARACTERISTICS:
 LENGTH: 33 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-961-083-227

Query Match 1.2%; Score 23; DB 3; Length 33;
 Best Local Similarity 100.0%; Pred. No. 0.033;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 TAAATCTACGACATAAAATC 23
 |||||||

Db 11 TAAATCTACGACATATAAATC 33

RESULT 6

US-08-731-716-1

; Sequence 1, Application US/08731716
; Patent No. 5789202

GENERAL INFORMATION:

APPLICANT: Hoskins, Joann

APPLICANT: Jaskunas, S. Richard

APPLICANT: Rostock, Pamela K.

APPLICANT: Zhao, Genshi

APPLICANT: Rostock, Paul R. Jr.

APPLICANT: No. 5789202's, Franklin H.

TITLE OF INVENTION: Penicillin Binding Protein From

TITLE OF INVENTION: Streptococcus Pneumoniae

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center

CITY: Indianapolis

STATE: Indiana

COUNTRY: U.S.

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/731,716

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Webster, Thomas D.

REGISTRATION NUMBER: 39,872

REFERENCE/DOCKET NUMBER: X-10,887

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3334

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2193 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHEICAL: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 1..2193

US-08-731-716-1

Query Match

Best Local Similarity 1.0%; Score 19; DB 1; Length 2193;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 TACGCTGCCTTGCAATG 1319

Db 1555 TACGCTGCCTTGCAATG 1573

RESULT 7

US-08-731-716-3

; Sequence 3, Application US/08731716
; Patent No. 5789202

GENERAL INFORMATION:

APPLICANT: Hoskins, Joann

APPLICANT: Jaskunas, S. Richard

APPLICANT: Rostock, Pamela K.

APPLICANT: Zhao, Genshi

APPLICANT: Rostock, Paul R. Jr.

APPLICANT: No. 5789202's, Franklin H.

TITLE OF INVENTION: Penicillin Binding Protein From

TITLE OF INVENTION: Streptococcus Pneumoniae

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center

CITY: Indianapolis

STATE: Indiana

COUNTRY: U.S.

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/731,716

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Webster, Thomas D.

REGISTRATION NUMBER: 39,872

REFERENCE/DOCKET NUMBER: X-10,887

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3334

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 2193 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: mRNA

HYPOTHEICAL: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 1..2193

US-08-731-716-3

Query Match

Best Local Similarity 1.0%; Score 19; DB 1; Length 2193;

Matches 13; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1301 TACGCTGCCTTGCAATG 1319

Db 1555 TACGCTGCCTTGCAATG 1573

RESULT 8

US-09-103-875-2/c

; Sequence 2, Application US/09103875A
; Patent No. 6221849

GENERAL INFORMATION:

APPLICANT: Szyt, Moshe

APPLICANT: Bigey, Pascal

APPLICANT: Ramchandani, Shyam

TITLE OF INVENTION: DNA METHYLTRANSFERASE GENOMIC SEQUENCES AND ANTISENSE

TITLE OF INVENTION: OLIGONUCLEOTIDES

FILE REFERENCE: 106101.194

CURRENT APPLICATION NUMBER: US/09/103,875A

CURRENT FILING DATE: 1998-06-24

EARLIER APPLICATION NUMBER: 60/069,865

EARLIER FILING DATE: 1997-12-17

EARLIER APPLICATION NUMBER: 08/866,340

EARLIER FILING DATE: 1997-05-30

NUMBER OF SEQ ID NOS: 138

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 2

LENGTH: 4136

TYPE: DNA

ORGANISM: Homo sapiens

US-09-103-875-2

Query Match 0.9%; Score 18; DB 4; Length 4136;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1589 CTATTGCTGCTATACG 1606
DB 3382 CTATTGCTGCTATACG 3365

RESULT 9

US-08-961-083-228/C
Sequence 228, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 228:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-083-228

Query Match 0.9%; Score 17; DB 3; Length 27;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1983 CTCACACGACACCA 1999
DB 27 CTCACACGACACCA 11

RESULT 10

US-09-020-956-5/C
Sequence 5, Application US/09020956
Patent No. 6261562
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillin, Davin C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 178
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,956
FILING DATE: 09-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 834 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-020-956-5

Query Match 0.9%; Score 17; DB 4; Length 834;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 AATTAATCACTCAT 30
DB 516 AATTAATCACTCAT 500

RESULT 11

US-09-030-607-5/C
Sequence 5, Application US/09030607
Patent No. 6262245
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillin, Davin C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
NUMBER OF SEQUENCES: 224
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,607
FILING DATE: 25-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 834 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-030-607-5

Query Match 0.9%; Score 17; DB 4; Length 834;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AATAAAATCAACTCAT 30
|||||
DB 516 AATAAAATCAACTCAT 500

RESULT 12

US-09-439-313-5/C
Sequence 5, Application US/09439313
Patent No. 6329505
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqun
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retler, Mark
APPLICANT: Solt, John
APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439.313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 834
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(834)
OTHER INFORMATION: n = A,T,C or G
US-09-439-313-5

Query Match 0.9%; Score 17; DB 4; Length 834;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AATAAAATCAACTCAT 30
|||||
DB 516 AATAAAATCAACTCAT 500

RESULT 13

US-08-630-915A-197/C
Sequence 197, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLER, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
TITLE OF INVENTION: USING SAME

NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630.915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741

INFORMATION FOR SEQ ID NO: 197:
SEQUENCE CHARACTERISTICS:
LENGTH: 971 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-630-915A-197

Query Match 0.9%; Score 17; DB 4; Length 971;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 AGGCTCAGGAGCTTGG 298
|||||
DB 373 AGGCTCAGGAGCTTGG 357

RESULT 14

US-09-444-053-3/C
Sequence 3, Application US/09444053A
Patent No. 6165728
GENERAL INFORMATION:
APPLICANT: Donna T. Ward
APPLICANT: Lex M. Cowser
TITLE OF INVENTION: ANTISENSE MODULATION OF NCK-2 EXPRESSION
FILE REFERENCE: RTS-0122
CURRENT APPLICATION NUMBER: US/09/444.053A
CURRENT FILING DATE: 1999-11-19
NUMBER OF SEQ ID NOS: 89
SEQ ID NO 3
LENGTH: 1457
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (133)..(1275)
US-09-444-053-3

Query Match 0.9%; Score 17; DB 4; Length 1457;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 AGGCTCAGGAGCTTGG 298
|||||
DB 677 AGGCTCAGGAGCTTGG 661

RESULT 15
US-09-444-053-55
Sequence 55, Application US/09444053A
Patent No. 6165728
GENERAL INFORMATION:
APPLICANT: Donna T. Ward
APPLICANT: Lex M. Cowsett
TITLE OF INVENTION: ANTISENSE MODULATION OF NCK-2 EXPRESSION
FILE REFERENCE: RTS-0122
CURRENT APPLICATION NUMBER: US/09/444, 053A
CURRENT FILING DATE: 1999-11-19
NUMBER OF SEQ ID NOS: 89
SEQ ID NO 55
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antisense Oligonucleotide
US-09-444-053-55

Query Match 0.8%; Score 16; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 282 AGGCTCAGGAAGCTTG 297
|||||
Db 5 aggcgcaggaagcttg 20

RESULT 16
US-09-457-708-21
Sequence 21, Application US/09457708
Patent No. 6326483
GENERAL INFORMATION:
APPLICANT: Kwiatkowski, David J.
APPLICANT: Sampson, Julian R.
APPLICANT: Povey, Sue
APPLICANT: van Slegtenhorst, Marjon
APPLICANT: Halley, Dicky
TITLE OF INVENTION: Compositions and Methods Based U
TITLE OF INVENTION: Sclerosis-1 (TSC-1) Gene and Gene Product
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Vinson & Elkins
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20004-1008
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/457, 708
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sanzo, Michael A.
REGISTRATION NUMBER: 36,912
REFERENCE/DOCKET NUMBER: BRI31/42002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 639-6585
TELEFAX: (202) 639-6604
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "PCR primer"
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-457-708-21

Query Match 0.8%; Score 16; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1713 ACCTGCTGAAGAAG 1728
|||||
Db 1 ACCTGCTGAAGAAG 16

RESULT 17
US-08-232-144-2/c
Sequence 2, Application US/08232144
Patent No. 5571695
GENERAL INFORMATION:
APPLICANT: SELBIE, Lisa
APPLICANT: HERZOG, Herbert
APPLICANT: SHINE, John
TITLE OF INVENTION: Human Neuropeptide Y-Y1 Receptor
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rothwell, Flagg, Ernst & Kurz
STREET: 555 13th St., N.W., Suite 701-East
CITY: Washington
STATE: DC
COUNTRY: US
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,144
FILING DATE: 26-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, Barbara G
REGISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 1871-107A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-783-6040
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-232-144-2

Query Match 0.8%; Score 16; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1765 CAGAATGGAATTC 1780
|||||
Db 27 CAGAATGGAATTC 12

RESULT 18
US-08-230-002-17
Sequence 17, Application US/08230002

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,080
FILING DATE: 17-JUL-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/592,126
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: INTRON 4 OF RAD50 GENOMIC SEQUENCE
US-08-687-080-65

Query Match 0.8%: Score 16; DB 2; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 612 ATGGACTACAAAGCTT 627
|||||
DB 33 ATGGACTACAAAGCTT 18

RESULT 21
US-08-592-126-137
Sequence 137, Application US/08592126
Patent No. 5821091
GENERAL INFORMATION:
APPLICANT: Gregory Doljanov
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
TITLE OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 151
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,126
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 137:
SEQUENCE CHARACTERISTICS:
LENGTH: 574 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: G115c.seq
US-08-592-126-137

Query Match 0.8%: Score 16; DB 1; Length 574;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1361 TTATGATGAGGAGTG 1376
|||||
DB 342 TTATGATGAGGAGTG 357

RESULT 22
US-08-784-289-1/c
Sequence 1, Application US/08784289
Patent No. 5817912
GENERAL INFORMATION:
APPLICANT: Pedrazzini, Thierry
APPLICANT: Brunner, Hans R.
TITLE OF INVENTION: Transgenic Animals with Disrupted Npy Y1
TITLE OF INVENTION: Receptor Genes
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Vinson & Elkins L.L.P.
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20004-1008
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,289
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sanzo, Michael A.
REGISTRATION NUMBER: 36,912
REFERENCE/DOCKET NUMBER: BMR 350/13000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 639-6585
TELEFAX: (202) 639-6604
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-784-289-1

Query Match 0.8%: Score 16; DB 1; Length 595;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmli, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 875 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-936-165A-125.

Query Match 0.8%; Score 16; DB 4; Length 875;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1064 GTTTACTAGTGAGTA 1079
DB 480 GTTTACTAGTGAGTA 495

RESULT 26
US-08-592-126-78
Sequence 78, Application US/08592126
Patent No. 5821091
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
NUMBER OF SEQUENCES: 151
CORRESPONDENCE ADDRESS:
ADDRESSEE: Denlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,126
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Shultz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 890 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: G256.seq
US-08-592-126-78

Query Match 0.8%; Score 16; DB 1; Length 890;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1556 ATCAAGACCTCTCAAT 1571
DB 310 ATCAAGACCTCTCAAT 325

RESULT 27
US-08-368-236-2
Sequence 2, Application US/08368236
Patent No. 5804408
GENERAL INFORMATION:
APPLICANT: Hagihara, et al.
TITLE OF INVENTION: A METHOD FOR EXPRESSING POLYPEPTIDES
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sherman and Shalloway
STREET: 413 N. Washington Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MD STORAGE
COMPUTER: IBM
OPERATING SYSTEM: DOS 5.1
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/368,236
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/941,139
FILING DATE: NO. 5804408ember 13, 1992
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Richard Steinberg.
REGISTRATION NUMBER: 26,588
REFERENCE/DOCKET NUMBER: S-2336
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 549-2282
TELEFAX: (703) 836-0106
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1186
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: NO
US-08-368-236-2

Query Match 0.8%; Score 16; DB 1; Length 1186;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 CGTTCTATCGAAGAC 118
DB 708 CGTTCTATCGAAGAC 723

RESULT 28
US-08-245-295-6/C
; Sequence 6, Application US/08245295
; Patent No. 5700658
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Kilgannon, Patrick D.
; TITLE OF INVENTION: ICAM-4 Materials and Methods
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,295
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27866/32055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1295 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-245-295-6

Query Match 0.8%; Score 16; DB 1; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTC 206
|||||
DB 723 AATTCCTCCAGGTC 708

RESULT 29
US-08-481-130-6/C
; Sequence 6, Application US/08481130
; Patent No. 5702917
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Kilgannon, Patrick D.
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,130
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32713
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1295 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-481-130-6

Query Match 0.8%; Score 16; DB 1; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTC 206
|||||
DB 723 AATTCCTCCAGGTC 708

RESULT 30
US-08-656-984A-6/C
; Sequence 6, Application US/08656984A
; Patent No. 5753502
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Kilgannon, Patrick D.
; TITLE OF INVENTION: ICAM-4 Materials and Methods

NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/656,984A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,604
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/33321
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ. ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1295 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-656-984A-6

Query Match 0.8%; Score 16; DB 1; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTG 206
|||||
Db 723 AATTCCTCCAGGTG 708

RESULT 31
US-08-485-604-6/C
Sequence 6, Application US/08485604
Patent No. 5773293
GENERAL INFORMATION:
APPLICANT: W.P. W. Michael
APPLICANT: Kilgannon, Patrick D.
TITLE OF INVENTION: ICAM-4 Materials and Methods

NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,604
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32715
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ. ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1295 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-485-604-6

Query Match 0.8%; Score 16; DB 1; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTG 206
|||||
Db 723 AATTCCTCCAGGTG 708

RESULT 32
US-08-487-595-6/C
Sequence 6, Application US/08487595
Patent No. 5852170
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Kilgannon, Patrick D.
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,595
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32714
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1295 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-487-595-6

Query Match 0.8%; Score 16; DB 2; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTG 206
|||||
DB 723 AATTCCTCCAGGTG 708

RESULT 33
US-08-863-790-25/c
Sequence 25, Application US/08863790
Patent No. 6087130
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Vazeux, Rosemay
TITLE OF INVENTION: ICAM-Related Protein
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Toole, Gerstein, Murray &
ADDRESSER: Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago

STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/863,790
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 6087130and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31570
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312)474-6300
TELEFAX: (312)474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1295 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-863-790-25

Query Match 0.8%; Score 16; DB 3; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTG 206
|||||
DB 723 AATTCCTCCAGGTG 708

RESULT 34
US-08-296-749-25/c
Sequence 25, Application US/08296749
Patent No. 6153395
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Vazeux, Rosemay
TITLE OF INVENTION: ICAM-Related Protein
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Toole, Gerstein, Murray &
ADDRESSER: Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/08/296,749
3 FILING DATE:
4 CLASSIFICATION: 435
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER: US 07/827,689
7 FILING DATE: 27-JAN-1992
8 PRIOR APPLICATION DATA:
9 APPLICATION NUMBER: US 07/889,724
10 FILING DATE: 26-MAY-1992
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: US 07/894,061
13 FILING DATE: 05-JUN-1992
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: US 08/009,266
16 FILING DATE: 22-JAN-1993
17 ATTORNEY/AGENT INFORMATION:
18 NAME: No. 615339sand, Greta E.
19 REGISTRATION NUMBER: 35,302
20 REFERENCE/DOCKET NUMBER: 31570
21 TELECOMMUNICATION INFORMATION:
22 TELEPHONE: (312)474-6300
23 TELEFAX: (312)474-0448
24 TELEX: 25-3856
25 INFORMATION FOR SEQ ID NO: 25:
26 SEQUENCE CHARACTERISTICS:
27 LENGTH: 1295 base pairs
28 TYPE: nucleic acid
29 STRANDEDNESS: single
30 TOPOLOGY: linear
31 MOLECULE TYPE: cDNA
32 US-08-296-749-25

Query Match 0.8%; Score 16; DB 3; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 191 AATTCCTCCAGGTG 206
DB 723 AATTCCTCCAGGTG 708

RESULT 35
US-08-440-845D-9
Sequence 9, Application US/08440845D
Patent No. 5955329
GENERAL INFORMATION:
APPLICANT: Yuan, L.
APPLICANT: Kridl, J.
APPLICANT: Dehesh, K.
APPLICANT: Knauf, V.
TITLE OF INVENTION: Engineering Plant Thioesterases for
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESS: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.40 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,845D
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Donna E. Scherer

1 REGISTRATION NUMBER: 34,719
2 NAME: Carl J. Schwedler
3 REGISTRATION NUMBER: 36,924
4 REFERENCE/DOCKET NUMBER: CGNE 113
5 TELECOMMUNICATION INFORMATION:
6 TELEPHONE: (916) 753-6313
7 TELEFAX: (916) 753-1510
8 INFORMATION FOR SEQ ID NO: 9:
9 SEQUENCE CHARACTERISTICS:
10 LENGTH: 1314 base pairs
11 TYPE: nucleic acid
12 STRANDEDNESS: single
13 TOPOLOGY: linear
14 MOLECULE TYPE: cDNA to mRNA
15 US-08-440-845D-9

Query Match 0.8%; Score 16; DB 2; Length 1314;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1364 AGTGATGGAGTGAAA 1379
DB 1058 AGTGATGGAGTGAAA 1073

RESULT 36
US-08-868-458-9
Sequence 9, Application US/08868458
Patent No. 6150512
GENERAL INFORMATION:
APPLICANT: Yuan, L.
TITLE OF INVENTION: Engineering Plant Thioesterases And
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESS: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
COMPUTER: IBM PC
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: Microsoft Word for Window 95 7.0a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/868,458
FILING DATE: 03-Jun-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/07064
FILING DATE: 15-MAY-96
APPLICATION NUMBER: 08/537,083
FILING DATE: 29-SEPT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 133
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1314 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-08-868-458-9

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-001-078A-2

Query Match 0.8%; Score 16; DB 2; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 663 ATTACTCAAGAGAGT 678
|||||
DB 494 ATTACTCAAGAGAGT 509

RESULT 40

US-08-463-218-2
; Sequence 2, Application US/08463218
; Patent No. 5986052
; GENERAL INFORMATION:
; APPLICANT: Goetlinck, Paul F.
; APPLICANT: Tondravi, Menrad
; APPLICANT: Binette, Francois
; TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,218
; FILING DATE: 05-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,078
; FILING DATE: 06-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: MGP-008DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-463-218-2

Query Match 0.8%; Score 16; DB 2; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 663 ATTACTCAAGAGAGT 678
|||||
DB 494 ATTACTCAAGAGAGT 509

RESULT 41

PCT-US94-00253-2
; Sequence 2, Application PC/US9400253
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX
; NUMBER OF SEQUENCES: 2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00253
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,078
; FILING DATE: 06-JAN-1993
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US94-00253-2

Query Match 0.8%; Score 16; DB 5; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 663 ATTACTCAAGAGAGT 678
|||||
DB 494 ATTACTCAAGAGAGT 509

RESULT 42

US-08-245-295-9/c
; Sequence 9, Application US/08245295
; Patent No. 5700658
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Kilgannon, Patrick D.
; TITLE OF INVENTION: ICAM-4 Materials and Methods
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,295
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27866/32055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1472 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-245-295-9

Query Match 0.8%; Score 16; DB 1; Length 1472;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATCCCTCCAGGTG 206
|||||
DB 723 AATCCCTCCAGGTG 708

RESULT 43
US-08-481-130-9/C
Sequence 9, Application US/08481130
Patent No. 5702917
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Kilgannon, Patrick D.
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,130
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:

NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32713
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1472 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-481-130-9

Query Match 0.8%; Score 16; DB 1; Length 1472;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATCCCTCCAGGTG 206
|||||
DB 723 AATCCCTCCAGGTG 708

RESULT 44
US-08-656-984A-9/C
Sequence 9, Application US/08656984A
Patent No. 5753502
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Kilgannon, Patrick D.
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/656,984A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,604
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:

NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/33321
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1472 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-656-984A-9

Query Match 0.8%; Score 16; DB 1; Length 1472;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTG 206
|||||
DB 723 AATTCCTCCAGGTG 708

RESULT 45
US-08-485-604-9/C
Sequence 9, Application US/08485604
Patent No. 5773293
GENERAL INFORMATION:
APPLICANT: W.P. Michael
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,604
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32715

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1472 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-485-604-9

Query Match 0.8%; Score 16; DB 1; Length 1472;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTG 206
|||||
DB 723 AATTCCTCCAGGTG 708

RESULT 46
US-08-487-595-9/C
Sequence 9, Application US/08487595
Patent No. 5852170
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,595
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32714
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448

TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1472 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-487-595-9

Query Match 0.8%; Score 16; DB 2; Length 1472;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATCCCTCCAGGTC 206
|||||
DB 723 AATCCCTCCAGGTC 708

RESULT 47

US-08-443-568B-15/c
Sequence 15, Application US/08443568B
Patent No. 5759807
GENERAL INFORMATION:
APPLICANT: Breese, Tim
APPLICANT: Havenga, Kirk
APPLICANT: Rindersknecht, Ernst
APPLICANT: Vandien, Richard
APPLICANT: Daniel, Yansura
TITLE OF INVENTION: Process for Producing Relaxin
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/443,568B
FILING DATE: 22-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/080,354
FILING DATE: 21-JUNE-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Abrams, Samuel B.
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 7842-037
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1500 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 438..1235
US-08-443-568B-15

Query Match 0.8%; Score 16; DB 1; Length 1500;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 874 GCTAGGAGCAGCCAT 889
|||||
DB 1474 GCTAGGAGCAGCCAT 1459

RESULT 48

PCT-US94-06997-15/c
Sequence 15, Application PC/TUS9406997
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: 460 Point San Bruno Boulevard
APPLICANT: South San Francisco, California
APPLICANT: United States of America
TITLE OF INVENTION: Process for Producing Relaxin
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06997
FILING DATE: 20-JUN-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: 7842-025-228
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-3660
TELEFAX: 415-854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1500 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 438..1235
PCT-US94-06997-15
Query Match 0.8%; Score 16; DB 5; Length 1500;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 874 GCTAGGAGCAGCCAT 889
|||||
DB 1474 GCTAGGAGCAGCCAT 1459

RESULT 49

US-07-759-568-4
Sequence 4, Application US/07759568
Patent No. 5374506
GENERAL INFORMATION:
APPLICANT: Murphy, Phillip M.
TITLE OF INVENTION: Cloning of cDNA Encoding a Functional
Human Interleukin-8 Receptor

NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cushman, Darby & Cushman
STREET: 1615 L Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/759,568
FILING DATE: 19910913
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Scott, Watson T.
REGISTRATION NUMBER: 26581
REFERENCE/DOCKET NUMBER: WTS/5683/91535/WBH
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 cush
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1510 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-07-759-568-4

Query Match 0.8%; Score 16; DB 1; Length 1510;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 432 ATTTAAGTTACCTCA 447
Db 7 ATTTAAGTTACCTCA 22

RESULT 50
US-08-288-899-1/C
Sequence 1, Application US/08288899
Patent No. 5610036
GENERAL INFORMATION:
APPLICANT: MURA, MASAMI
APPLICANT: ISHIDA, YUTAKA
APPLICANT: OI, HIDEYUKI
APPLICANT: MURAKAMI, YUKIMITSU
TITLE OF INVENTION: MUTANT AOX2 PROMOTER, MICROORGANISM
TITLE OF INVENTION: CARRYING SAME, METHOD OF PREPARATION THEREOF, AND
TITLE OF INVENTION: PRODUCTION OF HETEROLOGOUS PROTEIN USING SUCH
TITLE OF INVENTION: MICROORGANISM
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn Macpeak and Seas
STREET: 2100 Pennsylvania Avenue, NW
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/288,899

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/858,830
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Biggart, Waddell A.
REGISTRATION NUMBER: 24,861
REFERENCE/DOCKET NUMBER: 028202
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: plasmid DNA
FEATURE:
NAME/KEY: promoter
LOCATION: 1
US-08-288-899-1

Query Match 0.8%; Score 16; DB 1; Length 1528;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 811 AGACGATGATTCGCA 826
Db 931 AGACGATGATTCGCA 916

Search completed: June 13, 2002, 12:36:10
Job time: 10738 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 11:00:47 : Search time 3439.62 Seconds
(without alignments)
12573.529 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATCTACGACATATAAA.....ATCCTGACGACCAACCA 1999

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 21979536 seqs, 10817449327 residues

Word size : 15

Total number of hits satisfying chosen parameters: 41001

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/2/pna/PCRUS_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US095_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US096_COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US097_COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US098_COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US099_COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US100_COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US101_COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US102_COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US103_COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US104_COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US105_COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US106_COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US107_COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US108_COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US109_COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US110_COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US111_COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US112_COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US113_COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US114_COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US115_COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US116_COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US117_COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US118_COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US119_COMB.seq:*

44: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US6010_COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US6011_COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US6012_COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US6013_COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US6014_COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US6015_COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US6016_COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US6017_COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US6018_COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US6019_COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US6020_COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US6021_COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US6022_COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US6023_COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US6024_COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US6025_COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US6026_COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US6027_COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US6028_COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US6029_COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US6030_COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US6031_COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US6032_COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US6033_COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US6034_COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US6035_COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US6036_COMB.seq:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	1999	20	US-09-536-784-1
2	1999	100.0	1999	30	US-09-765-271-1
3	1999	100.0	1999	30	US-09-765-272-1
4	1999	100.0	10383	41	US-60-029-960-81
5	1999	100.0	10711	13	US-08-961-527-145
6	1642	82.1	2160	22	US-09-583-110-1312
7	1642	82.1	2166	15	US-09-107-433-1102
8	1438	71.9	2160	31	PCT-US02-03987-9325
9	1438	71.9	2160	31	US-09-815-242-8325
10	1438	71.9	2160	37	US-10-072-851-9325
11	1438	71.9	10333	45	US-60-061-998-996
12	796	39.8	960	5	US-08-116-541-3
13	541	27.1	2172	13	PCT-US97-14436-139
14	541	27.1	2172	13	US-08-911-503-139
15	541	27.1	2172	13	US-08-911-503A-139
16	293	14.7	9845	45	US-60-068-175-596
17	223	11.2	323	41	US-60-029-960-934
18	204	10.2	204	12	US-09-583-110-1313
19	25	1.3	8395	14	US-09-134-000-1534
20	25	1.3	8395	14	US-09-070-927-217
21	25	1.3	8395	14	US-09-070-927A-217
22	23	1.2	33	20	US-09-536-784-227
23	23	1.2	33	30	US-09-765-271-227
24	23	1.2	33	30	US-09-765-272-227
25	22	1.1	933	43	US-60-045-649-108
26	22	1.1	933	43	US-60-046-653-183
27	22	1.1	2337	31	PCT-US02-03987-6631
28	22	1.1	2337	31	US-09-815-242-6631
29	22	1.1	2337	37	US-10-072-851-6631
30	22	1.1	3478	45	US-60-068-217-710
31	21	1.1	537	25	US-09-649-163-2312

C	32	21	1.1	537	25	US-09-652-916-1790	Sequence 1790, App	105	20	1.0	32768	62	US-60-233-937-152	Sequence 152, App	
C	33	20	1.0	20	18	US-09-403-609A-25	Sequence 25, App1	C	106	20	1.0	41240	24	US-09-620-392-14275	Sequence 14275, App
C	34	20	1.0	242	20	US-09-535-896-44152	Sequence 44152, A	C	107	20	1.0	165972	28	US-09-702-134-9705	Sequence 9705, App
C	35	20	1.0	332	19	US-09-515-128-12893	Sequence 12893, A	C	108	20	1.0	165972	31	US-09-815-264-65753	Sequence 65753, A
C	36	20	1.0	355	10	US-08-688-870-855	Sequence 855, App	C	109	19	1.0	19	18	US-09-403-609-11	Sequence 11, App1
C	37	20	1.0	355	39	US-60-001-754-855	Sequence 855, App	C	110	19	1.0	19	18	US-09-403-609-12	Sequence 12, App1
C	38	20	1.0	371	16	US-09-205-070-44712	Sequence 44712, A	C	111	19	1.0	19	18	US-09-403-609A-11	Sequence 11, App1
C	39	20	1.0	371	17	US-09-321-214-31801	Sequence 31801, A	C	112	19	1.0	19	18	US-09-403-609A-12	Sequence 12, App1
C	40	20	1.0	371	17	US-09-340-623-44712	Sequence 44712, A	C	113	19	1.0	19	18	US-09-403-609A-24	Sequence 24, App1
C	41	20	1.0	371	19	US-09-516-335-31801	Sequence 31801, A	C	114	19	1.0	122	16	US-09-270-849B-11274	Sequence 61274, A
C	42	20	1.0	371	29	US-09-733-811-31801	Sequence 31801, A	C	115	19	1.0	156	25	US-09-654-617-381076	Sequence 381076, A
C	43	20	1.0	371	29	US-09-733-811A-31801	Sequence 31801, A	C	116	19	1.0	156	27	US-09-684-016-381076	Sequence 381076, A
C	44	20	1.0	371	33	US-09-898-888-44712	Sequence 44712, A	C	117	19	1.0	168	35	US-09-938-842A-447	Sequence 447, App
C	45	20	1.0	371	33	US-09-898-888A-44712	Sequence 44712, A	C	118	19	1.0	176	17	US-09-303-031A-279	Sequence 279, App
C	46	20	1.0	371	36	US-09-975-640-31801	Sequence 31801, A	C	119	19	1.0	176	33	US-09-894-949-279	Sequence 279, App
C	47	20	1.0	371	36	US-09-975-640A-31801	Sequence 31801, A	C	120	19	1.0	176	33	US-09-894-949A-279	Sequence 279, App
C	48	20	1.0	378	16	US-09-271-490-13695	Sequence 13695, A	C	121	19	1.0	245	17	US-09-303-031A-7415	Sequence 7415, App
C	49	20	1.0	378	34	US-09-925-552-13695	Sequence 13695, A	C	122	19	1.0	245	33	US-09-894-949-7415	Sequence 7415, App
C	50	20	1.0	378	37	US-10-032-354-13695	Sequence 13695, A	C	123	19	1.0	245	33	US-09-894-949A-7415	Sequence 7415, App
C	51	20	1.0	402	22	US-09-572-409-56717	Sequence 56717, A	C	124	19	1.0	276	17	US-09-304-517A-163891	Sequence 163891, A
C	52	20	1.0	415	17	US-09-359-067-47781	Sequence 47781, A	C	125	19	1.0	276	17	US-09-371-146A-163891	Sequence 163891, A
C	53	20	1.0	427	16	US-09-205-070-31561	Sequence 31561, A	C	126	19	1.0	276	36	US-09-985-678-163891	Sequence 163891, A
C	54	20	1.0	427	17	US-09-340-623-31561	Sequence 31561, A	C	127	19	1.0	286	17	US-09-304-517A-175873	Sequence 175873, A
C	55	20	1.0	427	33	US-09-898-888-31561	Sequence 31561, A	C	128	19	1.0	286	17	US-09-306-349A-26073	Sequence 26073, A
C	56	20	1.0	427	33	US-09-898-888A-31561	Sequence 31561, A	C	129	19	1.0	286	17	US-09-371-146A-175873	Sequence 175873, A
C	57	20	1.0	432	17	US-09-399-720-15214	Sequence 15214, A	C	130	19	1.0	286	36	US-09-960-481-26073	Sequence 26073, A
C	58	20	1.0	432	34	US-09-921-378-15214	Sequence 15214, A	C	131	19	1.0	286	36	US-09-985-678-175873	Sequence 175873, A
C	59	20	1.0	452	17	US-09-359-067-20940	Sequence 20940, A	C	132	19	1.0	291	17	US-09-303-031A-3625	Sequence 3625, App
C	60	20	1.0	469	19	US-09-528-409-113801	Sequence 113801, A	C	133	19	1.0	291	33	US-09-894-949-3625	Sequence 3625, App
C	61	20	1.0	469	35	US-09-933-524-113801	Sequence 113801, A	C	134	19	1.0	291	33	US-09-894-949A-3625	Sequence 3625, App
C	62	20	1.0	469	35	US-09-933-524A-113801	Sequence 113801, A	C	135	19	1.0	323	22	US-09-565-309A-20887	Sequence 20887, A
C	63	20	1.0	472	51	US-60-125-816-921	Sequence 921, App	C	136	19	1.0	381	23	US-09-606-977-54826	Sequence 54826, A
C	64	20	1.0	480	55	US-60-164-285-2087	Sequence 2087, App	C	137	19	1.0	393	29	US-09-724-866-23566	Sequence 23566, A
C	65	20	1.0	493	25	US-09-652-121-3163	Sequence 3163, App	C	138	19	1.0	393	29	US-09-724-866A-23566	Sequence 23566, A
C	66	20	1.0	501	1	PCT-US01-30732-3874	Sequence 3874, App	C	139	19	1.0	393	56	US-60-171-432-23566	Sequence 23566, A
C	67	20	1.0	502	24	US-09-634-306B-240647	Sequence 240647, App	C	140	19	1.0	406	23	US-09-606-977-54045	Sequence 54045, A
C	68	20	1.0	502	24	US-09-634-306B-240648	Sequence 240648, App	C	141	19	1.0	414	57	US-60-185-216-384	Sequence 384, App
C	69	20	1.0	541	37	US-10-020-338-20135	Sequence 20135, A	C	142	19	1.0	419	16	US-09-287-618-26734	Sequence 26734, A
C	70	20	1.0	585	24	US-09-634-306B-214358	Sequence 214358, A	C	143	19	1.0	428	22	US-09-565-309A-37361	Sequence 37361, A
C	71	20	1.0	680	66	US-60-278-232-260	Sequence 260, App	C	144	19	1.0	442	22	US-09-565-309A-37361	Sequence 37361, A
C	72	20	1.0	1299	16	US-09-252-691-2463	Sequence 2463, App	C	145	19	1.0	505	22	US-09-565-309A-20888	Sequence 20888, A
C	73	20	1.0	1299	16	US-09-252-691C-2463	Sequence 2463, App	C	146	19	1.0	507	24	US-09-634-306B-98352	Sequence 98352, A
C	74	20	1.0	1371	16	US-09-252-691-2450	Sequence 2450, App	C	147	19	1.0	537	32	US-09-849-526A-14660	Sequence 14660, A
C	75	20	1.0	1371	16	US-09-252-691C-2450	Sequence 2450, App	C	148	19	1.0	537	32	US-60-202-214-14415	Sequence 14415, A
C	76	20	1.0	1501	31	US-09-815-264-30030	Sequence 30030, App	C	149	19	1.0	599	33	US-09-865-439A-77247	Sequence 77247, A
C	77	20	1.0	1502	75	US-60-360-207-14183	Sequence 14183, A	C	150	19	1.0	599	33	US-09-865-439A-77247	Sequence 77247, A
C	78	20	1.0	1586	24	US-09-620-392-58865	Sequence 58865, A	C	151	19	1.0	647	61	US-60-207-458-121389	Sequence 121389, A
C	79	20	1.0	3475	75	US-60-360-207-9730	Sequence 9730, App	C	152	19	1.0	702	25	US-09-654-617-21939	Sequence 21939, App
C	80	20	1.0	5149	60	US-60-213-360-7352	Sequence 7352, App	C	153	19	1.0	702	27	US-09-684-016-21939	Sequence 21939, App
C	81	20	1.0	5149	66	US-60-278-258-15357	Sequence 15357, App	C	154	19	1.0	702	29	US-09-733-089-11409	Sequence 11409, A
C	82	20	1.0	5149	71	US-60-324-185-32088	Sequence 32088, A	C	155	19	1.0	702	31	US-09-816-660-11409	Sequence 11409, A
C	83	20	1.0	5314	27	US-09-698-010-13409	Sequence 13409, A	C	156	19	1.0	702	33	US-09-861-478-9784	Sequence 9784, App
C	84	20	1.0	5314	27	US-09-698-013-6642	Sequence 6642, App	C	157	19	1.0	748	33	US-09-861-478-1818	Sequence 1818, App
C	85	20	1.0	5314	27	US-09-698-014-5151	Sequence 5151, App	C	158	19	1.0	785	33	US-09-861-478-7204	Sequence 7204, App
C	86	20	1.0	5314	28	US-09-710-281-4646	Sequence 4646, App	C	159	19	1.0	810	61	US-60-229-524-171	Sequence 171, App
C	87	20	1.0	5314	29	US-09-726-788-6432	Sequence 6432, App	C	160	19	1.0	822	1	PCT-US02-30987-6970	Sequence 6970, App
C	88	20	1.0	5410	1	PCT-US01-10472-191	Sequence 191, App	C	161	19	1.0	822	31	US-09-815-247-6970	Sequence 6970, App
C	89	20	1.0	5766	56	US-60-172-373-12411	Sequence 12411, A	C	162	19	1.0	822	37	US-10-072-851-6970	Sequence 6970, App
C	90	20	1.0	5766	18	US-09-488-725A-155	Sequence 155, App	C	163	19	1.0	999	21	US-09-543-681A-1388	Sequence 1388, App
C	91	20	1.0	5766	22	US-09-598-042A-156	Sequence 156, App	C	164	19	1.0	1079	24	US-09-634-306B-250569	Sequence 250569, A
C	92	20	1.0	5766	37	US-10-098-841-156	Sequence 156, App	C	165	19	1.0	1079	24	US-09-634-306B-250570	Sequence 250570, A
C	93	20	1.0	6394	26	US-09-663-779-1130	Sequence 1130, App	C	166	19	1.0	1079	24	US-09-634-306B	

c 178	19	1.0	2251	27	US-09-684-016-271436	Sequence 271436,	251	18	0.9	25	26	US-09-660-220-7344	Sequence 7344, Ap
c 179	19	1.0	3715	24	US-09-620-392-35912	Sequence 35912, A	c 252	18	0.9	25	25	US-09-954-427-40492	Sequence 40492, A
c 180	19	1.0	3790	52	US-60-138-103-9711	Sequence 9711, Ap	c 253	18	0.9	25	62	US-09-233-166-40492	Sequence 40492, A
c 181	19	1.0	4299	45	US-60-068-175-461	Sequence 461, Ap	254	18	0.9	25	74	US-60-353-987-120805	Sequence 120805, A
c 182	19	1.0	5205	45	US-60-061-998-531	Sequence 531, Ap	255	18	0.9	69	16	US-09-270-767-28986	Sequence 28986, A
c 183	19	1.0	5448	1	PCT-US97-22578-29	Sequence 29, Appl	256	18	0.9	69	16	US-09-270-8498-9896	Sequence 9896, Ap
c 184	19	1.0	6720	24	US-09-620-392-44623	Sequence 44623, A	257	18	0.9	69	16	US-09-270-8498-115298	Sequence 115298, A
c 185	19	1.0	7679	18	US-09-404-520-5899	Sequence 5899, A	258	18	0.9	69	16	US-09-270-8498-115299	Sequence 115299, A
c 186	19	1.0	9678	24	US-09-620-392-47702	Sequence 47702, A	259	18	0.9	69	16	US-09-270-8498-115300	Sequence 115300, A
c 187	19	1.0	10181	28	US-09-702-134-8347	Sequence 8347, Ap	260	18	0.9	69	16	US-09-270-8498-115301	Sequence 115301, A
c 188	19	1.0	10181	31	US-09-815-264-61121	Sequence 61121, A	261	18	0.9	69	16	US-09-270-8498-115302	Sequence 115302, A
c 189	19	1.0	12692	33	US-09-897-516-2557	Sequence 2557, Ap	262	18	0.9	69	16	US-09-270-8498-115303	Sequence 115303, A
c 190	19	1.0	12692	33	US-09-897-516-2558	Sequence 2558, Ap	263	18	0.9	69	16	US-09-270-8498-115304	Sequence 115304, A
c 191	19	1.0	12692	33	US-09-897-516-2559	Sequence 2559, Ap	264	18	0.9	69	16	US-09-270-8498-115305	Sequence 115305, A
c 192	19	1.0	12692	33	US-09-897-516-2560	Sequence 2560, Ap	265	18	0.9	69	16	US-09-270-8498-115306	Sequence 115306, A
c 193	19	1.0	12692	33	US-09-897-516-2561	Sequence 2561, Ap	266	18	0.9	69	16	US-09-270-8498-115307	Sequence 115307, A
c 194	19	1.0	12692	33	US-09-897-516-2562	Sequence 2562, Ap	c 267	18	0.9	90	13	US-08-943-980-1334	Sequence 1334, Ap
c 195	19	1.0	12692	33	US-09-897-516-2563	Sequence 2563, Ap	c 268	18	0.9	90	41	US-60-027-838-1334	Sequence 1334, Ap
c 196	19	1.0	12692	33	US-09-897-516-2564	Sequence 2564, Ap	c 269	18	0.9	122	24	US-09-637-007-11955	Sequence 11955, A
c 197	19	1.0	12692	33	US-09-897-516-2565	Sequence 2565, Ap	270	18	0.9	125	13	US-08-992-332-1904	Sequence 1904, Ap
c 198	19	1.0	12692	33	US-09-897-516-2566	Sequence 2566, Ap	271	18	0.9	125	21	US-09-540-213-3859	Sequence 3859, Ap
c 199	19	1.0	12692	60	US-60-215-161-2557	Sequence 2557, Ap	272	18	0.9	125	42	US-60-033-364-452	Sequence 452, Ap
c 200	19	1.0	12692	60	US-60-215-161-2558	Sequence 2558, Ap	273	18	0.9	166	14	US-09-065-511-1939	Sequence 1939, Ap
c 201	19	1.0	12692	60	US-60-215-161-2559	Sequence 2559, Ap	274	18	0.9	178	17	US-09-313-294A-6301	Sequence 6301, Ap
c 202	19	1.0	12692	60	US-60-215-161-2560	Sequence 2560, Ap	275	18	0.9	178	47	US-08-086-722-6301	Sequence 6301, Ap
c 203	19	1.0	12692	60	US-60-215-161-2561	Sequence 2561, Ap	276	18	0.9	192	20	US-09-532-315-25590	Sequence 25590, A
c 204	19	1.0	12692	60	US-60-215-161-2562	Sequence 2562, Ap	277	18	0.9	192	42	US-60-039-051-951	Sequence 951, Ap
c 205	19	1.0	12692	60	US-60-215-161-2563	Sequence 2563, Ap	278	18	0.9	205	8	US-08-404-891-1272	Sequence 1272, Ap
c 206	19	1.0	12692	60	US-60-215-161-2564	Sequence 2564, Ap	279	18	0.9	205	20	US-09-539-801-610	Sequence 610, Ap
c 207	19	1.0	12692	60	US-60-215-161-2565	Sequence 2565, Ap	280	18	0.9	207	20	US-09-534-843-18795	Sequence 18795, A
c 208	19	1.0	12692	60	US-60-215-161-2566	Sequence 2566, Ap	281	18	0.9	207	41	US-60-023-278-509	Sequence 509, Ap
c 209	19	1.0	19466	61	US-60-229-524-37	Sequence 37, Appl	282	18	0.9	207	57	US-60-184-770-1292	Sequence 1292, Ap
c 210	19	1.0	19504	59	US-60-207-352-25	Sequence 25, Appl	283	18	0.9	211	13	US-08-900-106-3256	Sequence 3256, Ap
c 211	19	1.0	19651	24	US-09-630-392-35883	Sequence 35883, A	284	18	0.9	211	13	US-08-998-886-3908	Sequence 3908, Ap
c 212	19	1.0	19651	28	US-09-702-134-5081	Sequence 5081, Ap	285	18	0.9	211	20	US-09-539-806-24590	Sequence 24590, Ap
c 213	19	1.0	19651	31	US-09-815-264-63556	Sequence 63556, A	286	18	0.9	211	21	US-09-540-208-97214	Sequence 97214, A
c 214	19	1.0	20034	60	US-60-212-656-258	Sequence 258, Ap	287	18	0.9	213	16	US-09-270-8498-164425	Sequence 164425, A
c 215	19	1.0	20085	62	US-60-230-435-457	Sequence 457, Ap	288	18	0.9	214	11	US-08-725-693-3723	Sequence 3723, A
c 216	19	1.0	22377	63	US-60-245-221-68	Sequence 68, Appl	289	18	0.9	214	20	US-09-539-8334-18330	Sequence 18330, A
c 217	19	1.0	23041	1	PCT-US01-01341-1753	Sequence 1753, Ap	c 290	18	0.9	216	17	US-09-313-922A-2424	Sequence 2424, Ap
c 218	19	1.0	23041	30	US-09-764-864-1753	Sequence 1753, Ap	c 291	18	0.9	216	47	US-60-086-722-424	Sequence 424, Ap
c 219	19	1.0	23041	37	US-10-080-129-1753	Sequence 1753, Ap	292	18	0.9	222	13	US-08-918-671-859	Sequence 859, Ap
c 220	19	1.0	23135	1	PCT-US01-01341-1754	Sequence 1754, Ap	293	18	0.9	222	21	US-09-540-029-35977	Sequence 35977, A
c 221	19	1.0	23135	30	US-09-764-864-1754	Sequence 35892, A	294	18	0.9	222	41	US-60-026-899-859	Sequence 859, Ap
c 222	19	1.0	23135	30	US-09-764-864-1754	Sequence 1754, Ap	295	18	0.9	227	17	US-09-333-617-99902	Sequence 99902, A
c 223	19	1.0	23135	30	US-09-764-905-35892	Sequence 35892, A	296	18	0.9	227	25	US-09-654-617-99902	Sequence 99902, A
c 224	19	1.0	23135	37	US-10-080-129-1754	Sequence 1754, Ap	297	18	0.9	227	27	US-09-684-016-99902	Sequence 99902, A
c 225	19	1.0	23135	37	US-10-092-399-35892	Sequence 35892, A	298	18	0.9	227	33	US-09-878-978-13569	Sequence 13569, A
c 226	19	1.0	23260	24	US-09-620-392-5504	Sequence 5504, Ap	299	18	0.9	229	13	US-08-943-978-3370	Sequence 3370, Ap
c 227	19	1.0	23260	28	US-09-702-134-25269	Sequence 25269, A	300	18	0.9	229	20	US-09-539-8008-2484	Sequence 2484, Ap
c 228	19	1.0	32768	59	US-09-815-264-64698	Sequence 64698, A	301	18	0.9	229	20	US-09-539-8008-2484	Sequence 2484, Ap
c 229	19	1.0	32768	62	US-60-207-583-162	Sequence 162, Ap	302	18	0.9	230	13	US-08-923-993-2446	Sequence 2446, Ap
c 230	19	1.0	32768	62	US-60-230-445-229	Sequence 229, Ap	303	18	0.9	230	21	US-09-540-208-28018	Sequence 28018, A
c 231	19	1.0	58144	64	US-60-258-275-164	Sequence 164, Ap	304	18	0.9	230	41	US-60-025-467-2446	Sequence 2446, Ap
c 232	19	1.0	62318	60	US-60-242-679-400	Sequence 400, Ap	305	18	0.9	231	11	US-08-725-693-3723	Sequence 3723, A
c 233	19	1.0	69198	63	US-60-245-225-209	Sequence 209, Ap	c 306	18	0.9	231	13	US-08-962-919-2593	Sequence 2593, Ap
c 234	19	1.0	102312	20	US-09-534-859-350	Sequence 350, Ap	307	18	0.9	231	20	US-09-539-8334-15261	Sequence 15261, A
c 235	19	1.0	102312	31	US-09-803-736-350	Sequence 350, Ap	c 308	18	0.9	231	21	US-09-540-733-1121	Sequence 1121, Ap
c 236	19	1.0	106412	19	US-08-528-237A-1802	Sequence 1802, Ap	c 309	18	0.9	232	24	US-09-637-007-11722	Sequence 11722, A
c 237	19	1.0	176723	60	US-60-212-664-244	Sequence 244, Ap	310	18	0.9	236	14	US-09-041-885-378	Sequence 378, Ap
c 238	19	1.0	228454	19	US-09-528-237A-1439	Sequence 1439, Ap	311	18	0.9	236	20	US-09-534-847-18310	Sequence 18310, A
c 239	19	1.0	1830121	8	US-08-426-787-1	Sequence 1, Appl	312	18	0.9	236	41	US-60-022-912-378	Sequence 378, Ap
c 240	19	1.0	1830121	8	US-08-426-787-1	Sequence 1, Appl	313	18	0.9	238	8	US-08-404-891-203	Sequence 203, Ap
c 241	19	1.0	1830121	8	US-08-487-429-1	Sequence 1, Appl	314	18	0.9	238	20	US-09-539-801-1708	Sequence 1708, Ap
c 242	19	1.0	1830121	8	US-08-487-429-1	Sequence 1, Appl	315	18	0.9	238	24	US-09-637-007-11707	Sequence 11707, Ap
c 243	19	1.0	1830121	22	US-09-557-884-1	Sequence 1, Appl	316	18	0.9	245	23	US-09-606-977-3218	Sequence 3218, Ap
c 244	19	1.0	1830121	22	US-09-557-884-1	Sequence 1, Appl	317	18	0.9	246	29	US-09-724-866-24092	Sequence 24092, A
c 245	19	1.0	1830121	25	US-09-643-990A-1	Sequence 1, Appl	318	18	0.9	246	29	US-09-724-866A-24092	Sequence 24092, A
c 246	19	1.0	1830121	25	US-09-643-990A-1	Sequence 1, Appl	319	18	0.9	246	56	US-60-171-312-24092	Sequence 24092, A
c 247	18	0.9	1830121	25	US-09-643-990A-1	Sequence 9, Appl	320	18	0.9	250	14	US-09-035-172-8249	Sequence 8249, Ap
c 248	18	0.9	18	18	US-09-403-609-9	Sequence 10, Appl	321	18	0.9	250	24	US-09-637-007-11666	Sequence 11666, A
c 249	18	0.9	18	18	US-09-403-609A-9	Sequence 9, Appl	322	18	0.9	253	5	US-08-196-365-8252	Sequence 8252, Ap
c 250	18	0.9	18	18	US-09-403-609A-10	Sequence 10, Appl	323	18	0.9	253	5	US-08-196-363A-8252	Sequence 8252, Ap

324	18	0.9	253	5	US-08-196-363-8252	Sequence 8252, Ap	397	18	0.9	295	13	US-08-918-972-519	Sequence 519, App
325	18	0.9	253	32	US-09-859-490-8252	Sequence 8253, Ap	C 398	18	0.9	295	19	US-09-521-660-5720	Sequence 5720, Ap
326	18	0.9	253	20	US-09-534-856-22314	Sequence 22314, A	399	18	0.9	295	21	US-09-540-744-13456	Sequence 13456, A
327	18	0.9	255	20	US-09-539-800-15921	Sequence 15921, A	400	18	0.9	295	41	US-60-025-153-519	Sequence 519, App
328	18	0.9	255	20	US-09-539-800B-15921	Sequence 15921, A	401	18	0.9	306	64	US-60-253-655-24789	Sequence 24789, A
329	18	0.9	255	49	US-60-100-256-675	Sequence 675, App	402	18	0.9	306	64	US-60-255-559-24789	Sequence 24789, A
330	18	0.9	256	18	US-09-420-691-542	Sequence 542, App	403	18	0.9	315	19	US-09-521-640-100128	Sequence 100128, A
331	18	0.9	256	20	US-09-534-856-22356	Sequence 22356, A	C 404	18	0.9	315	17	US-09-304-517A-90675	Sequence 90675, A
C 332	18	0.9	257	5	US-08-196-363-1304	Sequence 1304, Ap	C 405	18	0.9	318	17	US-09-371-146A-90675	Sequence 90675, A
C 333	18	0.9	257	5	US-08-196-363A-1304	Sequence 1304, Ap	C 406	18	0.9	318	22	US-09-553-099-792	Sequence 292, App
C 334	18	0.9	257	5	US-08-196-363-1304	Sequence 1304, Ap	C 407	18	0.9	318	25	US-09-654-617-266151	Sequence 266151, A
C 335	18	0.9	257	13	US-08-978-620-4656	Sequence 4656, Ap	C 408	18	0.9	318	27	US-09-684-016-266151	Sequence 266151, A
C 336	18	0.9	257	20	US-09-532-315-25586	Sequence 25586, A	C 409	18	0.9	318	36	US-09-985-678-90675	Sequence 90675, A
C 337	18	0.9	257	32	US-09-859-490-1304	Sequence 1304, Ap	C 410	18	0.9	318	52	US-60-130-189-4012	Sequence 4012, Ap
C 338	18	0.9	258	64	US-60-253-457-2819	Sequence 2819, Ap	C 411	18	0.9	318	64	US-60-253-654-23030	Sequence 23030, A
C 339	18	0.9	260	14	US-09-073-079-930	Sequence 930, App	412	18	0.9	328	64	US-60-255-559-23030	Sequence 23030, A
C 340	18	0.9	260	21	US-09-540-229-59569	Sequence 59569, A	413	18	0.9	328	64	US-60-255-559-23030	Sequence 23030, A
C 341	18	0.9	260	43	US-60-047-802-930	Sequence 930, App	414	18	0.9	328	64	US-60-255-559-23030	Sequence 23030, A
C 342	18	0.9	264	20	US-09-539-800-16898	Sequence 16898, A	415	18	0.9	328	64	US-60-255-559-23030	Sequence 23030, A
C 343	18	0.9	264	20	US-09-539-800B-16898	Sequence 16898, A	416	18	0.9	328	64	US-60-255-559-23030	Sequence 23030, A
C 344	18	0.9	265	18	US-09-421-106-32113	Sequence 32113, A	417	18	0.9	337	33	US-09-985-439A-92993	Sequence 92993, A
C 345	18	0.9	265	19	US-09-521-640-300868	Sequence 300868, A	418	18	0.9	337	59	US-60-207-458-137008	Sequence 137008, A
C 346	18	0.9	267	12	US-08-878-507-1120	Sequence 1120, Ap	419	18	0.9	340	64	US-60-253-654-23022	Sequence 23022, A
C 347	18	0.9	267	13	US-08-940-864-2116	Sequence 2116, Ap	420	18	0.9	342	17	US-09-324-828-643	Sequence 828, Ap
C 348	18	0.9	267	14	US-09-044-767-3921	Sequence 3921, Ap	421	18	0.9	342	17	US-09-998-388-643	Sequence 643, App
C 349	18	0.9	267	20	US-09-539-806-15087	Sequence 15087, A	422	18	0.9	342	47	US-60-087-831-643	Sequence 831, App
C 350	18	0.9	267	21	US-09-540-212A-11871	Sequence 11871, A	423	18	0.9	343	16	US-09-271-490-4586	Sequence 4586, Ap
C 351	18	0.9	267	21	US-09-540-764-35442	Sequence 35442, A	424	18	0.9	343	17	US-09-321-21A-8846	Sequence 8846, Ap
C 352	18	0.9	267	21	US-60-016-092-1120	Sequence 1120, Ap	425	18	0.9	343	19	US-09-516-335-8846	Sequence 8846, Ap
C 353	18	0.9	267	41	US-60-027-236-2116	Sequence 2116, Ap	426	18	0.9	343	29	US-09-733-811-8846	Sequence 8846, Ap
C 354	18	0.9	268	14	US-09-041-894-2542	Sequence 2542, Ap	427	18	0.9	343	29	US-09-733-811-8846	Sequence 8846, Ap
C 355	18	0.9	268	20	US-09-539-806-39339	Sequence 39339, A	428	18	0.9	343	34	US-09-925-552-4586	Sequence 4586, Ap
C 356	18	0.9	268	43	US-60-040-199-2542	Sequence 2542, Ap	429	18	0.9	343	36	US-09-975-640A-8846	Sequence 8846, Ap
C 357	18	0.9	270	13	US-08-986-693A-3916	Sequence 3916, Ap	430	18	0.9	343	36	US-09-975-640A-8846	Sequence 8846, Ap
C 358	18	0.9	270	13	US-08-986-693A-3916	Sequence 3916, Ap	431	18	0.9	343	37	US-10-032-334-4586	Sequence 4586, Ap
C 359	18	0.9	271	21	US-09-540-229-76687	Sequence 76687, A	C 432	18	0.9	348	19	US-09-521-640-159119	Sequence 159119, A
C 360	18	0.9	271	21	US-09-540-212A-55512	Sequence 55512, A	C 433	18	0.9	349	25	US-09-654-617-257180	Sequence 257180, A
C 361	18	0.9	272	12	US-08-822-285-3561	Sequence 3561, Ap	434	18	0.9	349	27	US-09-684-016-257180	Sequence 257180, A
C 362	18	0.9	272	12	US-09-539-806-33158	Sequence 33158, A	C 435	18	0.9	352	52	US-60-132-884-616	Sequence 884, App
C 363	18	0.9	272	17	US-60-013-696-3561	Sequence 3561, Ap	436	18	0.9	356	64	US-60-253-654-14501	Sequence 14501, A
C 364	18	0.9	274	17	US-09-370-505-1359	Sequence 1359, Ap	437	18	0.9	358	24	US-08-668-226-419	Sequence 419, App
C 365	18	0.9	274	20	US-09-534-852-621	Sequence 621, App	438	18	0.9	358	21	US-09-540-222-81619	Sequence 81619, A
C 366	18	0.9	274	48	US-60-096-463-1359	Sequence 1359, Ap	439	18	0.9	359	16	US-09-271-122-5383	Sequence 5383, Ap
C 367	18	0.9	275	13	US-08-923-903-7935	Sequence 735, App	440	18	0.9	359	28	US-09-713-841-5383	Sequence 5383, Ap
C 368	18	0.9	275	17	US-09-304-517A-54251	Sequence 54251, A	441	18	0.9	359	34	US-09-927-875A-5383	Sequence 5383, Ap
C 369	18	0.9	275	17	US-09-371-146A-54251	Sequence 54251, A	442	18	0.9	359	34	US-09-927-875A-5383	Sequence 5383, Ap
C 370	18	0.9	275	21	US-09-540-208-30032	Sequence 30032, A	C 443	18	0.9	359	24	US-09-637-007-11501	Sequence 11501, A
C 371	18	0.9	275	25	US-09-654-617-304653	Sequence 304653, A	444	18	0.9	370	18	US-09-410-515-47	Sequence 47, App1
C 372	18	0.9	275	27	US-09-684-016-304653	Sequence 304653, A	445	18	0.9	370	31	US-09-823-241-8057	Sequence 8057, Ap
C 373	18	0.9	275	36	US-09-985-678-54251	Sequence 54251, A	446	18	0.9	370	32	US-09-836-377-47	Sequence 47, App1
C 374	18	0.9	275	41	US-60-025-467-7935	Sequence 735, App	447	18	0.9	370	32	US-09-836-377-47	Sequence 47, App1
C 375	18	0.9	276	11	US-08-725-693-2129	Sequence 2129, Ap	C 448	18	0.9	371	24	US-09-637-007-11496	Sequence 11496, A
C 376	18	0.9	276	20	US-09-539-633-11321	Sequence 11321, A	C 449	18	0.9	371	24	US-09-637-007-11496	Sequence 11496, A
C 377	18	0.9	277	17	US-09-359-922-7641	Sequence 7641, Ap	C 450	18	0.9	373	16	US-09-270-848B-171024	Sequence 171024, A
C 378	18	0.9	277	17	US-09-359-922-7641	Sequence 7641, Ap	451	18	0.9	374	16	US-09-248-797-17222	Sequence 17222, A
C 379	18	0.9	277	21	US-09-540-229-123720	Sequence 123720, A	452	18	0.9	374	34	US-09-925-556-17222	Sequence 17222, A
C 380	18	0.9	278	21	US-08-961-524-796	Sequence 796, App	453	18	0.9	375	16	US-09-248-797-17319	Sequence 17319, A
C 381	18	0.9	278	21	US-09-540-208-29359	Sequence 29359, A	454	18	0.9	375	34	US-09-925-556-17319	Sequence 17319, A
C 382	18	0.9	278	21	US-09-540-766-44339	Sequence 44339, A	C 455	18	0.9	376	24	US-09-637-007-11483	Sequence 11483, A
C 383	18	0.9	279	12	US-08-878-507-72	Sequence 72, App1	C 456	18	0.9	378	17	US-09-359-067-19833	Sequence 19833, A
C 384	18	0.9	279	21	US-09-540-212A-26668	Sequence 26668, A	C 457	18	0.9	379	17	US-09-362-510-53654	Sequence 53654, A
C 385	18	0.9	279	40	US-60-016-092-72	Sequence 72, App1	C 458	18	0.9	379	34	US-09-362-510A-53654	Sequence 53654, A
C 386	18	0.9	282	10	US-08-668-236-1083	Sequence 1083, Ap	C 459	18	0.9	379	34	US-09-904-01A-53654	Sequence 53654, A
C 387	18	0.9	282	21	US-09-540-229-64426	Sequence 64426, A	460	18	0.9	380	24	US-09-637-068A-21179	Sequence 21179, A
C 388	18	0.9	283	13	US-08-918-671-1163	Sequence 1163, Ap	461	18	0.9	380	24	US-09-637-068B-21179	Sequence 21179, A
C 389	18	0.9	283	16	US-09-229-413B-2892	Sequence 2892, Ap	462	18	0.9	380	25	US-09-654-617-214097	Sequence 214097, A
C 390	18	0.9	283	21	US-09-540-229-58465	Sequence 58465, A	463	18	0.9	380	27	US-09-684-016-214097	Sequence 214097, A
C 391	18	0.9	283	21	US-09-540-766-53407	Sequence 53407, A	464	18	0.9	381	16	US-09-235-076-36962	Sequence 36962, A
C 392	18	0.9	283	34	US-09-922-293-2892	Sequence 2892, Ap	465	18	0.9	381	17	US-09-272-722-14814	Sequence 14814, A
C 393	18	0.9	283	41	US-60-026-899-1163	Sequence 1163, Ap	466	18	0.9	381	17	US-09-332-788-36962	Sequence 36962, A
C 394	18	0.9	285	15	US-09-107-909-133	Sequence 133, App	467	18	0.9	381	29	US-09-737-223-36962	Sequence 36962, A
C 395	18	0.9	289	20	US-09-534-856-22311	Sequence 22311, A	468	18	0.9	381	34	US-09-909-621-14814	Sequence 14814, A
C 396	18	0.9	289	53	US-60-142-699-940	Sequence 940, App	469	18	0.9	382	17	US-09-359-067-6671	Sequence 6671, Ap

C 470	18	0.9	382	18	US-09-431-517-26265	Sequence 26265, A	C 543	18	0.9	438	17	US-09-399-720-18707	Sequence 18707, A
C 471	18	0.9	382	24	US-09-637-007-11195	Sequence 11195, A	C 544	18	0.9	438	19	US-09-521-640-94813	Sequence 94813, A
C 472	18	0.9	383	18	US-09-489-036-21023	Sequence 21023, A	C 545	18	0.9	438	34	US-09-904-809-2168	Sequence 2168, Ap
C 473	18	0.9	383	19	US-09-515-126-12004	Sequence 12004, A	C 546	18	0.9	438	34	US-09-904-939-28120	Sequence 28120, A
C 474	18	0.9	383	19	US-09-521-640-205375	Sequence 205375, A	C 547	18	0.9	438	34	US-09-921-378-18707	Sequence 18707, A
C 475	18	0.9	383	22	US-09-577-409-12004	Sequence 12004, A	C 548	18	0.9	438	34	US-09-925-554-44257	Sequence 44257, A
C 476	18	0.9	383	35	US-09-943-143-21023	Sequence 21023, A	C 439	18	0.9	439	1	PCT-US01-01302-183	Sequence 183, App
C 477	18	0.9	384	35	US-09-867-701-1947	Sequence 1947, Ap	C 550	18	0.9	439	1	PCT-US01-01309-365	Sequence 365, App
C 478	18	0.9	384	59	US-60-207-484-1947	Sequence 1947, Ap	C 551	18	0.9	439	30	US-09-764-681-183	Sequence 183, App
C 479	18	0.9	387	18	US-09-489-036-21714	Sequence 21714, A	C 552	18	0.9	443	19	US-09-521-660-22607	Sequence 22607, A
C 480	18	0.9	387	35	US-09-943-143-21714	Sequence 21714, A	C 553	18	0.9	443	19	US-09-528-409-82803	Sequence 82803, A
C 481	18	0.9	388	32	US-09-849-526A-28080	Sequence 28080, A	C 554	18	0.9	443	35	US-09-933-524-82803	Sequence 82803, A
C 482	18	0.9	388	59	US-60-202-214-27835	Sequence 27835, A	C 555	18	0.9	443	35	US-09-933-524A-82803	Sequence 82803, A
C 483	18	0.9	394	19	US-09-521-640-191470	Sequence 191470, A	C 556	18	0.9	445	36	US-09-998-598-504	Sequence 904, App
C 484	18	0.9	394	53	US-60-140-769-19121	Sequence 19121, A	C 557	18	0.9	446	22	US-09-565-309A-32887	Sequence 32887, A
C 485	18	0.9	397	17	US-09-388-907-1071	Sequence 1071, Ap	C 558	18	0.9	449	17	US-09-306-330A-13645	Sequence 13645, A
C 486	18	0.9	397	18	US-09-438-762A-20996	Sequence 20996, A	C 559	18	0.9	449	34	US-09-909-629-13645	Sequence 13645, A
C 487	18	0.9	400	23	US-09-606-977-43778	Sequence 43778, A	C 560	18	0.9	451	58	US-60-197-872-36252	Sequence 36252, A
C 488	18	0.9	401	16	US-09-271-122-7965	Sequence 7965, Ap	C 561	18	0.9	456	30	US-09-770-444-527	Sequence 527, App
C 489	18	0.9	401	28	US-09-713-841-7965	Sequence 7965, Ap	C 562	18	0.9	457	17	US-09-353-690-9175	Sequence 9175, App
C 490	18	0.9	401	28	US-09-927-875-7965	Sequence 7965, Ap	C 563	18	0.9	457	24	US-09-634-306B-94603	Sequence 94603, A
C 491	18	0.9	401	34	US-09-927-875A-7965	Sequence 7965, Ap	C 564	18	0.9	457	24	US-09-634-306B-94604	Sequence 94604, A
C 492	18	0.9	403	16	US-08-271-122-7966	Sequence 7966, Ap	C 565	18	0.9	457	24	US-09-634-306B-305498	Sequence 305498, A
C 493	18	0.9	403	17	US-09-359-067-23060	Sequence 23060, A	C 566	18	0.9	457	24	US-09-634-306B-305499	Sequence 305499, A
C 494	18	0.9	403	18	US-09-489-036-5423	Sequence 5423, Ap	C 567	18	0.9	457	24	US-09-634-306B-305500	Sequence 305500, A
C 495	18	0.9	403	28	US-09-713-841-7966	Sequence 7966, Ap	C 568	18	0.9	457	28	US-09-705-926-20766	Sequence 20766, A
C 496	18	0.9	403	34	US-09-927-875-7966	Sequence 7966, Ap	C 569	18	0.9	457	34	US-09-922-340A-9175	Sequence 9175, Ap
C 497	18	0.9	403	34	US-09-927-875A-7966	Sequence 7966, Ap	C 570	18	0.9	457	34	US-09-922-340A-9175	Sequence 9175, Ap
C 498	18	0.9	403	34	US-09-943-143-5423	Sequence 5423, Ap	C 571	18	0.9	458	19	US-09-521-640-216982	Sequence 216982, A
C 499	18	0.9	404	25	US-09-637-086A-51580	Sequence 51580, A	C 572	18	0.9	461	18	US-09-436-762A-22844	Sequence 22844, A
C 500	18	0.9	404	25	US-09-637-086D-51580	Sequence 51580, A	C 573	18	0.9	461	23	US-09-614-150-25283	Sequence 25283, A
C 501	18	0.9	407	17	US-09-388-906A-12836	Sequence 12836, A	C 574	18	0.9	461	58	US-60-191-617-25399	Sequence 25399, A
C 502	18	0.9	409	17	US-09-399-720-11528	Sequence 11528, A	C 575	18	0.9	461	58	US-60-191-661-252008	Sequence 252008, A
C 503	18	0.9	409	29	US-09-726-809-275	Sequence 275, App	C 576	18	0.9	466	26	US-09-666-335A-17711	Sequence 17711, Ap
C 504	18	0.9	409	34	US-09-921-378-11528	Sequence 11528, A	C 577	18	0.9	473	16	US-09-235-076-19144	Sequence 19144, A
C 505	18	0.9	410	16	US-08-293-972-30098	Sequence 30098, A	C 578	18	0.9	473	16	US-09-289-768-25459	Sequence 25459, A
C 506	18	0.9	410	16	US-09-465-231-1884	Sequence 1884, Ap	C 579	18	0.9	473	17	US-09-332-762-19144	Sequence 19144, A
C 507	18	0.9	410	36	US-09-904-939-30098	Sequence 30098, A	C 580	18	0.9	473	25	US-09-737-223-19144	Sequence 19144, A
C 508	18	0.9	410	36	US-09-983-965-1884	Sequence 1884, Ap	C 581	18	0.9	473	35	US-09-939-397-25459	Sequence 25459, A
C 509	18	0.9	414	18	US-09-496-911-6649	Sequence 6649, Ap	C 582	18	0.9	474	20	US-09-539-806-43266	Sequence 43266, A
C 510	18	0.9	414	36	US-09-975-673A-6649	Sequence 6649, Ap	C 583	18	0.9	474	25	US-09-644-165A-1394	Sequence 1394, Ap
C 511	18	0.9	415	24	US-08-637-086A-52413	Sequence 52413, A	C 584	18	0.9	476	23	US-09-606-977-11246	Sequence 11246, Ap
C 512	18	0.9	415	24	US-09-637-086D-52413	Sequence 52413, A	C 585	18	0.9	476	25	US-09-654-617-416587	Sequence 416587, A
C 513	18	0.9	416	16	US-09-235-076-16942	Sequence 16942, A	C 586	18	0.9	476	26	US-09-669-817A-25424	Sequence 25424, A
C 514	18	0.9	416	16	US-09-289-768-4703	Sequence 4703, Ap	C 587	18	0.9	476	27	US-09-684-016-116587	Sequence 116587, A
C 515	18	0.9	416	16	US-09-289-768-34928	Sequence 34928, Ap	C 588	18	0.9	477	23	US-09-606-977-19501	Sequence 19501, A
C 516	18	0.9	416	17	US-09-332-782-16942	Sequence 16942, A	C 589	18	0.9	478	24	US-09-634-306B-286371	Sequence 286371, A
C 517	18	0.9	416	29	US-09-737-223-16942	Sequence 16942, A	C 590	18	0.9	484	16	US-09-234-611-14008	Sequence 14008, A
C 518	18	0.9	416	35	US-09-939-397-4703	Sequence 4703, Ap	C 591	18	0.9	484	16	US-09-235-076-19170	Sequence 19170, A
C 519	18	0.9	416	35	US-09-939-397-4928	Sequence 34928, A	C 592	18	0.9	484	16	US-09-289-768-20951	Sequence 20951, A
C 520	18	0.9	417	16	US-09-248-797-2405	Sequence 2405, Ap	C 593	18	0.9	484	17	US-09-332-782-19170	Sequence 19170, A
C 521	18	0.9	417	17	US-09-333-534-1267	Sequence 1267, Ap	C 594	18	0.9	484	29	US-09-737-223-19170	Sequence 19170, A
C 522	18	0.9	417	34	US-09-925-564-2405	Sequence 2405, Ap	C 595	18	0.9	484	34	US-09-904-809-14008	Sequence 14008, A
C 523	18	0.9	418	25	US-09-654-617-199496	Sequence 199496, A	C 596	18	0.9	484	35	US-09-939-397-20951	Sequence 20951, A
C 524	18	0.9	418	25	US-09-684-016-199496	Sequence 199496, A	C 597	18	0.9	491	1	PCT-US01-08631-8417	Sequence 8417, Ap
C 525	18	0.9	423	57	US-60-197-872-32223	Sequence 32223, A	C 598	18	0.9	491	32	US-09-850-147-518	Sequence 518, App
C 526	18	0.9	424	24	US-09-637-007-6896	Sequence 6896, Ap	C 599	18	0.9	491	51	US-60-202-213-53	Sequence 53, App1
C 527	18	0.9	425	30	US-09-770-175-3740	Sequence 3740, Ap	C 600	18	0.9	498	22	US-09-587-373-53	Sequence 53, App1
C 528	18	0.9	425	32	US-09-834-366-38662	Sequence 38662, A	C 601	18	0.9	498	57	US-60-184-770-540	Sequence 540, App
C 529	18	0.9	425	58	US-60-197-873-38662	Sequence 38662, A	C 602	18	0.9	499	54	US-60-207-458-64708	Sequence 64708, A
C 530	18	0.9	426	19	US-09-521-640-207928	Sequence 207928, A	C 603	18	0.9	503	24	US-09-637-007-11016	Sequence 11016, A
C 531	18	0.9	431	18	US-09-465-231-2773	Sequence 2773, Ap	C 604	18	0.9	506	17	US-09-396-970-3208	Sequence 3208, Ap
C 532	18	0.9	431	36	US-09-983-965-2773	Sequence 2773, Ap	C 605	18	0.9	507	37	US-10-021-333-11980	Sequence 11980, Ap
C 533	18	0.9	434	17	US-09-332-782-10344	Sequence 10344, A	C 606	18	0.9	507	64	US-60-225-619-11980	Sequence 11980, A
C 534	18	0.9	434	19	US-09-515-694-10344	Sequence 10344, A	C 607	18	0.9	509	17	US-09-371-508-1457	Sequence 1457, Ap
C 535	18	0.9	435	19	US-08-528-409-43019	Sequence 43019, A	C 608	18	0.9	509	17	US-09-371-508-1457	Sequence 1457, Ap
C 536	18	0.9	435	35	US-09-933-524-43019	Sequence 43019, A	C 609	18	0.9	509	29	US-09-747-508-1457	Sequence 1457, Ap
C 537	18	0.9	435	35	US-09-933-524A-43019	Sequence 43019, A	C 610	18	0.9	512	29	US-09-758-472-2814	Sequence 2814, Ap
C 538	18	0.9	436	36	US-09-496-911-6650	Sequence 6650, Ap	C 611	18	0.9	512	31	US-09-815-343-1069	Sequence 1069, Ap
C 539	18	0.9	436	36	US-09-975-673A-6650	Sequence 6650, Ap	C 612	18	0.9	513	21	US-09-540-229-10284	Sequence 10284, Ap
C 540	18	0.9	438	16	US-09-234-611-2168	Sequence 2168, Ap	C 613	18	0.9	513	57	US-60-184-774-1265	Sequence 1265, Ap
C 541	18	0.9	438	16	US-09-248-797-44257	Sequence 44257, A	C 614	18	0.9	513	58	US-60-196-710-27	Sequence 27, App1
C 542	18	0.9	438	16	US-09-293-972-28120	Sequence 28120, A	C 615	18	0.9	515	21	US-09-540-764-57370	Sequence 57370, A

C 616	18	0.9	525	US-09-565-309A-32889	Sequence 32889, A	C 689	18	0.9	855	US-09-654-617-393981	Sequence 393981,
C 617	18	0.9	525	US-09-865-439A-115077	Sequence 115077,	C 690	18	0.9	855	US-09-684-016-393981	Sequence 393981,
C 618	18	0.9	532	US-09-634-306B-135305	Sequence 135305,	C 691	18	0.9	855	US-09-733-089-9790	Sequence 9790, Ap
C 619	18	0.9	532	US-09-637-007-10952	Sequence 10952, A	C 692	18	0.9	855	US-09-816-660-9790	Sequence 9790, Ap
C 620	18	0.9	538	US-09-654-617-132803	Sequence 132803,	C 693	18	0.9	870	US-09-248-796-2367	Sequence 2367, Ap
C 621	18	0.9	538	US-09-684-016-132803	Sequence 132803,	C 694	18	0.9	870	US-09-60-409-2367	Sequence 2367, Ap
C 622	18	0.9	541	PCT-US01-00663-11033	Sequence 11033, A	C 695	18	0.9	903	US-09-654-617-452299	Sequence 452299,
C 623	18	0.9	541	US-09-864-761-8173	Sequence 8173, Ap	C 696	18	0.9	903	US-09-684-016-452299	Sequence 452299,
C 624	18	0.9	542	US-09-540-229-97845	Sequence 97845, A	C 697	18	0.9	927	US-09-565-309A-67162	Sequence 67162, A
C 625	18	0.9	542	US-09-668-683-11706	Sequence 11706, A	C 698	18	0.9	932	US-09-526-597A-1	Sequence 1, ApP1
C 626	18	0.9	542	US-60-168-614-537	Sequence 537, App	C 699	18	0.9	932	US-09-574-735C-1	Sequence 1, ApP1
C 627	18	0.9	544	US-09-505-532-34714	Sequence 34714, A	C 700	18	0.9	962	US-60-173-464-15037	Sequence 15037, A
C 628	18	0.9	544	US-09-819-091A-34714	Sequence 34714, A	C 701	18	0.9	962	US-60-191-637-18322	Sequence 18322, A
C 629	18	0.9	552	US-09-565-309A-32891	Sequence 32891, A	C 702	18	0.9	962	US-60-191-681-14497	Sequence 14497, A
C 630	18	0.9	553	US-09-637-007-10925	Sequence 10925, A	C 703	18	0.9	963	US-09-654-617-24475	Sequence 24475, A
C 631	18	0.9	553	US-09-668-683-4370	Sequence 4370, Ap	C 704	18	0.9	963	US-09-684-016-24475	Sequence 24475, A
C 632	18	0.9	553	US-60-185-215-1431	Sequence 1431, Ap	C 705	18	0.9	963	US-09-733-089-11984	Sequence 11984, A
C 633	18	0.9	557	US-60-173-464-22682	Sequence 22682, A	C 706	18	0.9	963	US-09-816-660-11984	Sequence 11984, A
C 634	18	0.9	560	US-60-226-326-4833	Sequence 4833, Ap	C 707	18	0.9	965	US-60-167-217-18330	Sequence 18330, A
C 635	18	0.9	562	US-09-637-007-10912	Sequence 10912, A	C 708	18	0.9	972	US-09-248-796-7674	Sequence 7674, Ap
C 636	18	0.9	562	US-09-824-559-9529	Sequence 9529, Ap	C 709	18	0.9	972	US-60-096-409-7674	Sequence 7674, Ap
C 637	18	0.9	569	US-09-865-439A-85167	Sequence 85167, A	C 710	18	0.9	975	US-09-654-617-390366	Sequence 390366,
C 638	18	0.9	569	US-60-207-458-129309	Sequence 129309,	C 711	18	0.9	975	US-09-684-016-390366	Sequence 390366,
C 639	18	0.9	571	US-09-540-229-98054	Sequence 98054, A	C 712	18	0.9	987	US-09-667-495-941	Sequence 941, App
C 640	18	0.9	571	US-09-565-309A-35738	Sequence 35738, A	C 713	18	0.9	987	US-09-710-281-5701	Sequence 5701, Ap
C 641	18	0.9	571	US-09-565-309A-54894	Sequence 54894, A	C 714	18	0.9	987	US-09-659-999-7415	Sequence 7415, Ap
C 642	18	0.9	575	US-60-213-362-8174	Sequence 8174, Ap	C 715	18	0.9	987	US-09-716-475-7762	Sequence 7762, Ap
C 643	18	0.9	581	US-09-634-306B-81926	Sequence 81926, A	C 716	18	0.9	1075	US-09-654-617-259810	Sequence 259810,
C 644	18	0.9	593	US-09-873-402A-77121	Sequence 77121, A	C 717	18	0.9	1075	US-09-684-016-259810	Sequence 259810,
C 645	18	0.9	593	US-09-654-617-282226	Sequence 282226,	C 718	18	0.9	1104	PCT-US01-08631-3271	Sequence 3271, Ap
C 646	18	0.9	595	US-09-684-016-282226	Sequence 282226,	C 719	18	0.9	1136	US-09-705-926-2668	Sequence 2668, Ap
C 647	18	0.9	603	US-10-021-323-16236	Sequence 16236, A	C 720	18	0.9	1136	US-09-572-411-5886	Sequence 5886, Ap
C 648	18	0.9	603	US-60-255-619-16236	Sequence 16236, A	C 721	18	0.9	1149	US-60-360-207-28762	Sequence 28762, A
C 649	18	0.9	610	US-09-396-970-3257	Sequence 3257, Ap	C 722	18	0.9	1153	US-60-340-187-1137	Sequence 1137, Ap
C 650	18	0.9	610	US-09-637-886-7232	Sequence 7232, Ap	C 723	18	0.9	1197	US-09-565-309A-65030	Sequence 65030, A
C 651	18	0.9	617	US-60-279-526-480	Sequence 480, App	C 724	18	0.9	1197	US-09-595-332A-1033	Sequence 1033, App
C 652	18	0.9	631	US-60-171-494-692	Sequence 692, App	C 725	18	0.9	1200	PCT-US01-12230-65	Sequence 65, ApP1
C 653	18	0.9	634	US-09-565-309A-30337	Sequence 30337, A	C 726	18	0.9	1200	US-09-835-188-65	Sequence 85, ApP1
C 654	18	0.9	634	US-09-565-309A-32893	Sequence 32893, A	C 727	18	0.9	1216	US-09-770-175-8705	Sequence 8705, Ap
C 655	18	0.9	635	US-09-565-309A-48997	Sequence 48997, A	C 728	18	0.9	1229	PCT-US00-26524B-781	Sequence 781, App
C 656	18	0.9	635	US-09-758-463-364	Sequence 364, App	C 729	18	0.9	1241	US-09-565-309A-52670	Sequence 52670, A
C 657	18	0.9	639	US-60-127-461-2743	Sequence 2743, Ap	C 730	18	0.9	1245	US-09-205-070-15512	Sequence 15512, A
C 658	18	0.9	647	US-09-849-529A-4261	Sequence 4261, Ap	C 731	18	0.9	1245	US-09-340-623-15512	Sequence 15512, A
C 659	18	0.9	647	US-60-196-868-4261	Sequence 4261, Ap	C 732	18	0.9	1245	US-09-898-888-15512	Sequence 15512, A
C 660	18	0.9	651	US-09-565-309A-67164	Sequence 67164, A	C 733	18	0.9	1245	US-09-898-888-15512	Sequence 15512, A
C 661	18	0.9	653	US-60-253-457-29218	Sequence 29218, A	C 734	18	0.9	1247	US-60-126-590-446	Sequence 446, App
C 662	18	0.9	664	US-09-637-007-10784	Sequence 10784, A	C 735	18	0.9	1281	US-09-565-309A-52066	Sequence 52066, A
C 663	18	0.9	672	US-09-873-402A-90571	Sequence 90571, A	C 736	18	0.9	1281	US-09-565-309A-61884	Sequence 61884, A
C 664	18	0.9	672	US-60-209-830-62310	Sequence 62310, A	C 737	18	0.9	1281	US-09-595-329A-1968	Sequence 1968, Ap
C 665	18	0.9	674	US-60-192-746-653	Sequence 653, App	C 738	18	0.9	1299	PCT-US02-03987-6177	Sequence 6177, Ap
C 666	18	0.9	674	US-60-194-109-493	Sequence 493, App	C 739	18	0.9	1299	US-09-815-242-6177	Sequence 6177, Ap
C 667	18	0.9	677	US-10-060-063-807	Sequence 807, App	C 740	18	0.9	1328	US-10-072-851-6177	Sequence 6177, Ap
C 668	18	0.9	691	US-09-565-309A-30340	Sequence 30340, A	C 741	18	0.9	1328	US-09-359-922-13066	Sequence 13066, A
C 669	18	0.9	696	US-60-126-265-4104	Sequence 4104, Ap	C 742	18	0.9	1328	US-09-59-922-13066	Sequence 13066, A
C 670	18	0.9	699	US-09-505-532-47728	Sequence 47728, A	C 743	18	0.9	1360	US-09-270-767-13093	Sequence 13093, A
C 671	18	0.9	699	US-09-819-091A-47728	Sequence 47728, A	C 744	18	0.9	1360	US-09-471-275-10244	Sequence 10244, A
C 672	18	0.9	700	US-09-735-271-1227	Sequence 1227, Ap	C 745	18	0.9	1439	US-09-572-411-6001	Sequence 6001, App
C 673	18	0.9	711	US-09-333-534-10984	Sequence 10984, A	C 746	18	0.9	1439	US-08-716-345A-1	Sequence 1, ApP1
C 674	18	0.9	744	US-09-565-309A-67163	Sequence 67163, A	C 747	18	0.9	1501	US-09-815-264-793	Sequence 793, App
C 675	18	0.9	751	US-60-172-373-16289	Sequence 16289, A	C 748	18	0.9	1501	US-09-815-264-20538	Sequence 20538, A
C 676	18	0.9	751	US-60-184-747-16289	Sequence 16289, A	C 749	18	0.9	1501	US-09-815-264-22322	Sequence 22322, A
C 677	18	0.9	755	US-09-733-507-6	Sequence 6, ApP1	C 750	18	0.9	1531	US-09-471-275-5400	Sequence 5400, Ap
C 678	18	0.9	787	US-09-698-012-5715	Sequence 5715, Ap	C 751	18	0.9	1539	US-09-340-187-113	Sequence 113, App
C 679	18	0.9	816	PCT-US01-14827-4236	Sequence 4236, Ap	C 752	18	0.9	1555	US-09-634-306B-253001	Sequence 253001,
C 680	18	0.9	816	US-09-577-408-5155	Sequence 5155, Ap	C 753	18	0.9	1555	US-09-634-306B-253002	Sequence 253002,
C 681	18	0.9	816	US-60-340-187-823	Sequence 823, App	C 754	18	0.9	1555	US-09-634-306B-253003	Sequence 253003,
C 682	18	0.9	816	US-60-340-187-824	Sequence 824, App	C 755	18	0.9	1555	US-09-634-306B-253004	Sequence 253004,
C 683	18	0.9	850	US-09-404-520-20799	Sequence 20799, A	C 756	18	0.9	1564	US-60-184-770-1355	Sequence 1355, A
C 684	18	0.9	850	US-09-634-306B-157285	Sequence 157285,	C 757	18	0.9	1644	PCT-US99-22855-3048	Sequence 3048, Ap
C 685	18	0.9	851	US-09-205-070-2635	Sequence 2635, Ap	C 758	18	0.9	1666	US-09-665-489-97	Sequence 97, ApP1
C 686	18	0.9	851	US-09-340-623-2635	Sequence 2635, Ap	C 759	18	0.9	1772	PCT-US00-31919-2	Sequence 2, ApP1
C 687	18	0.9	851	US-09-898-888-2635	Sequence 2635, Ap	C 760	18	0.9	1772	US-09-862-419-2	Sequence 2, ApP1
C 688	18	0.9	851	US-09-898-888A-2635	Sequence 2635, Ap	C 761	18	0.9	1788	US-09-614-150-18266	Sequence 18266, A

762	18	0.9	1811	56	US-60-173-464-26808	Sequence 26808, A	c 835	18	0.9	5748	31	US-09-815-264-80609	Sequence 80609, A
c 763	18	0.9	1999	24	US-09-620-392-60322	Sequence 60322, A	836	18	0.9	6250	23	PCT-US01-08631-10864	Sequence 10864, A
c 764	18	0.9	2019	24	US-09-620-392-60321	Sequence 22318, A	c 837	18	0.9	6334	23	US-09-614-150-25282	Sequence 25282, A
c 765	18	0.9	2029	53	US-60-144-351-4513	Sequence 4513, Ap	c 838	18	0.9	6334	58	US-60-191-637-25398	Sequence 25398, A
766	18	0.9	2064	18	US-09-488-7254-1636	Sequence 1636, Ap	c 839	18	0.9	6334	58	US-60-191-661-50007	Sequence 20007, A
767	18	0.9	2064	25	US-09-653-450A-228	Sequence 228, Ap	840	18	0.9	6335	56	US-60-173-464-22681	Sequence 22681, A
768	18	0.9	2093	1	PCT-US99-08504-57	Sequence 57, Appl	c 841	18	0.9	7187	19	US-09-528-2378A-1100	Sequence 1100, Ap
769	18	0.9	2093	16	US-09-298-102-57	Sequence 57, Appl	c 842	18	0.9	7187	1	PCT-US01-08631-10440	Sequence 10440, Ap
c 770	18	0.9	2101	23	US-09-614-150-7319	Sequence 7319, Ap	c 843	18	0.9	8887	1	PCT-US01-08656-1004	Sequence 3004, Ap
c 771	18	0.9	2101	23	US-09-619-049-374	Sequence 374, Ap	c 844	18	0.9	8941	56	US-60-173-464-15036	Sequence 15036, A
c 772	18	0.9	2184	74	US-60-356-051-812	Sequence 812, Ap	c 845	18	0.9	8941	58	US-60-191-637-18321	Sequence 18321, A
c 773	18	0.9	2211	18	US-09-417-507-11293	Sequence 11293, A	c 846	18	0.9	8941	58	US-60-191-661-14496	Sequence 14496, A
774	18	0.9	2263	55	US-60-164-285-3478	Sequence 3478, Ap	c 847	18	0.9	8944	55	US-60-167-217-18329	Sequence 18329, A
775	18	0.9	2263	55	US-60-164-285-4186	Sequence 4186, Ap	c 848	18	0.9	9359	1	PCT-US01-08631-22954	Sequence 22954, A
776	18	0.9	2263	71	US-60-324-185-2081	Sequence 2081, Ap	c 849	18	0.9	9502	18	US-09-471-275-2917	Sequence 2917, Ap
777	18	0.9	2269	56	US-60-172-360-22628	Sequence 22628, A	c 850	18	0.9	9767	23	US-09-614-150-18265	Sequence 18265, A
778	18	0.9	2288	1	PCT-US99-13418-53	Sequence 53, Appl	c 851	18	0.9	10280	1	PCT-US01-01334-23187	Sequence 23187, A
779	18	0.9	2288	18	US-09-461-325-53	Sequence 53, Appl	c 852	18	0.9	10280	30	US-09-764-905-23187	Sequence 23187, A
780	18	0.9	2288	37	US-10-012-542-53	Sequence 53, Appl	c 853	18	0.9	10280	37	US-10-092-399-23187	Sequence 23187, A
781	18	0.9	2288	37	US-10-050-873-53	Sequence 53, Appl	c 854	18	0.9	10339	1	PCT-US01-01338-597	Sequence 597, Ap
782	18	0.9	2313	46	US-60-070-771-408	Sequence 408, Ap	c 855	18	0.9	10339	1	PCT-US01-01338-597	Sequence 597, Ap
783	18	0.9	2387	18	US-09-471-275-935	Sequence 935, Ap	c 856	18	0.9	10339	1	PCT-US01-01339-6747	Sequence 6747, Ap
784	18	0.9	2387	18	US-09-488-725A-5208	Sequence 5208, Ap	c 857	18	0.9	10339	30	US-09-764-881-6747	Sequence 6747, Ap
785	18	0.9	2404	25	US-09-644-871-9451	Sequence 9451, Ap	c 858	18	0.9	10339	37	US-10-091-572-597	Sequence 597, Ap
786	18	0.9	2404	25	US-09-644-873-11222	Sequence 11222, A	c 859	18	0.9	12376	54	US-60-150-582-64	Sequence 64, Appl
787	18	0.9	2404	25	US-09-649-162-9323	Sequence 9323, Ap	c 860	18	0.9	12666	63	US-60-243-468-327	Sequence 2927, Ap
788	18	0.9	2404	25	US-09-652-109-10083	Sequence 10083, A	861	18	0.9	13165	1	PCT-US01-01334-2924	Sequence 2924, Ap
789	18	0.9	2404	25	US-09-652-121-6845	Sequence 6845, Ap	862	18	0.9	13165	1	PCT-US01-01333-459	Sequence 459, Ap
790	18	0.9	2404	25	US-09-652-122-6711	Sequence 6711, Ap	863	18	0.9	13165	35	US-09-950-083-8868	Sequence 8868, Ap
791	18	0.9	2404	25	US-09-652-123-9756	Sequence 9756, Ap	864	18	0.9	13165	37	US-10-079-979-9294	Sequence 2924, Ap
792	18	0.9	2404	25	US-09-652-126-8944	Sequence 8944, Ap	c 865	18	0.9	15186	3	PCT-US01-01334-7948	Sequence 7948, Ap
793	18	0.9	2404	27	US-09-658-012-8739	Sequence 8739, Ap	c 866	18	0.9	15186	30	US-09-764-874-7948	Sequence 7948, Ap
800	18	0.9	2404	27	US-09-698-013-7250	Sequence 7250, Ap	c 873	18	0.9	17684	31	US-09-815-264-22046	Sequence 22046, A
801	18	0.9	2404	28	US-09-700-000-6114	Sequence 6114, Ap	874	18	0.9	17758	1	PCT-US01-01334-18950	Sequence 38950, A
802	18	0.9	2404	28	US-09-710-281-5151	Sequence 5151, Ap	875	18	0.9	17758	30	US-09-764-905-38950	Sequence 38950, A
803	18	0.9	2404	28	US-09-710-282-3345	Sequence 3345, Ap	876	18	0.9	17758	37	US-10-092-399-38950	Sequence 38950, A
804	18	0.9	2404	28	US-09-717-350-5797	Sequence 5797, Ap	c 877	18	0.9	20564	63	US-60-242-679-533	Sequence 533, Ap
805	18	0.9	2404	29	US-09-726-172-2261	Sequence 2261, Ap	c 878	18	0.9	22008	1	PCT-US01-01334-8170	Sequence 8170, Ap
806	18	0.9	2404	29	US-09-726-175-3037	Sequence 3037, Ap	879	18	0.9	22008	30	US-09-764-869-2110	Sequence 2110, Ap
807	18	0.9	2404	29	US-09-726-211-1565	Sequence 1565, Ap	880	18	0.9	22008	30	US-09-764-869-2110	Sequence 2110, Ap
c 808	18	0.9	2463	28	US-09-708-427-28468	Sequence 28468, A	c 881	18	0.9	22008	30	US-09-764-874-8170	Sequence 8170, Ap
c 809	18	0.9	2463	33	US-09-887-272A-198	Sequence 198, Ap	882	18	0.9	22008	35	US-09-950-082-1108	Sequence 1108, Ap
c 810	18	0.9	2463	35	US-09-938-842A-1341	Sequence 1341, Ap	883	18	0.9	22008	37	US-10-091-504-2110	Sequence 2110, Ap
c 811	18	0.9	2733	1	PCT-US01-08631-17918	Sequence 17918, A	c 884	18	0.9	22008	37	US-10-092-400-8170	Sequence 8170, Ap
c 812	18	0.9	2733	24	US-09-634-306B-11835	Sequence 11835, A	885	18	0.9	22336	19	US-09-528-2378A-2866	Sequence 2866, A
813	18	0.9	2804	52	US-60-138-103-5329	Sequence 5329, Ap	886	18	0.9	22336	14	US-09-620-392-24463	Sequence 24463, A
c 814	18	0.9	2887	18	US-09-404-520-25370	Sequence 25370, A	887	18	0.9	22639	28	US-09-453-702B-173	Sequence 173, Ap
c 815	18	0.9	2890	22	US-09-584-852-6989	Sequence 6989, Ap	c 887	18	0.9	22639	28	US-09-620-392-39156	Sequence 39156, A
816	18	0.9	2958	37	US-10-060-087-5524	Sequence 5524, Ap	c 888	18	0.9	23839	28	US-09-702-134-23976	Sequence 23976, A
c 817	18	0.9	3011	75	US-60-360-207-6548	Sequence 6548, Ap	c 889	18	0.9	25321	31	US-09-815-264-60471	Sequence 60471, A
818	18	0.9	3025	18	US-09-404-520-1128	Sequence 1128, Ap	c 890	18	0.9	25321	19	US-09-528-2378A-2866	Sequence 2866, A
c 819	18	0.9	3201	70	US-60-316-362-2190	Sequence 2190, Ap	891	18	0.9	27133	24	US-09-620-392-24463	Sequence 24463, A
c 820	18	0.9	3293	70	US-60-316-362-1190	Sequence 1190, Ap	892	18	0.9	28690	14	US-09-070-927A-138	Sequence 138, Ap
821	18	0.9	3327	23	US-09-614-150-24424	Sequence 24424, A	893	18	0.9	28690	14	US-09-070-927A-138	Sequence 138, Ap
822	18	0.9	3327	56	US-60-171-627-595	Sequence 595, Ap	894	18	0.9	29650	53	US-60-144-331-1946	Sequence 1946, Ap
823	18	0.9	3327	56	US-60-173-464-5960	Sequence 5960, Ap	c 895	18	0.9	32187	1	PCT-US01-01332-1550	Sequence 1550, Ap
824	18	0.9	3327	58	US-60-191-637-7339	Sequence 7339, Ap	c 896	18	0.9	32187	37	US-09-764-847-1550	Sequence 1550, Ap
825	18	0.9	3327	58	US-60-191-681-5728	Sequence 5728, Ap	c 897	18	0.9	32187	37	US-10-092-814-1550	Sequence 1550, Ap
826	18	0.9	3330	55	US-60-167-217-7435	Sequence 7435, Ap	c 898	18	0.9	32193	1	PCT-US01-01332-1549	Sequence 1549, Ap
c 827	18	0.9	3824	1	PCT-US01-08631-17884	Sequence 17884, A	c 899	18	0.9	32193	30	US-09-764-847-1549	Sequence 1549, Ap
828	18	0.9	4101	23	US-09-614-150-7318	Sequence 7318, Ap	c 900	18	0.9	32193	37	US-10-092-154-1549	Sequence 1549, Ap
829	18	0.9	4101	23	US-09-619-049-373	Sequence 373, Ap	c 901	18	0.9	32768	59	US-60-200-366-16	Sequence 16, Appl
c 830	18	0.9	4136	30	US-09-777-525-2	Sequence 2, Appl1	c 902	18	0.9	32768	59	US-60-200-366-17	Sequence 17, Appl
831	18	0.9	4994	1	PCT-US01-08631-24425	Sequence 24425, A	c 903	18	0.9	32768	59	US-60-200-366-18	Sequence 18, Appl
c 832	18	0.9	5169	24	US-09-620-392-49938	Sequence 49938, A	c 904	18	0.9	32768	59	US-60-200-366-19	Sequence 19, Appl
c 833	18	0.9	5293	70	US-60-316-362-190	Sequence 190, Ap	c 905	18	0.9	32768	62	US-60-230-445-66	Sequence 66, Appl
c 834	18	0.9	5748	28	US-09-702-134-10127	Sequence 10127, A	c 906	18	0.9	32768	62	US-60-230-445-71	Sequence 71, Appl
							907	18	0.9	32768	62	US-60-234-446-132	Sequence 132, Ap


```

Db 1 TAAATCTAGCAATAAAAATCACTATGCTGACTGGGTTCTGAAGCGGCTCA 60
QY 61 TGGCCAGGTAATGATATCCACAGATTTGGTTAAGGAATGCTTTATGAGAACCA 120
Db 61 TGGCCAGGTAATGATATCCACAGATTTGGTTAAGGAATGCTTTATGAGAACCA 120
QY 121 TGGCTTCTCGACACAGGGGGATTTGATACATCCGTATCTGGGAGCTTTCTGGCAA 180
Db 121 TGGCTTCTCGACACAGGGGGATTTGATACATCCGTATCTGGGAGCTTTCTGGCAA 180
QY 181 TGTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCACCAACAGTTGATTAAGTTGAC 240
Db 181 TGTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCACCAACAGTTGATTAAGTTGAC 240
QY 241 TTAAGTTTCACTGAGCTTCGACAGACTATTTCTGTAAGGCTCAGGAAGCTTGAT 300
Db 241 TTAAGTTTCACTGAGCTTCGACAGACTATTTCTGTAAGGCTCAGGAAGCTTGAT 300
QY 301 AGCGATTCTAGTAGAACAACAAGCAACAGCAAGAAATCTGACCTACTATATAATAA 360
Db 301 AGCGATTCTAGTAGAACAACAAGCAACAGCAAGAAATCTGACCTACTATATAATAA 360
QY 361 GGTCTACATGCTAATGGGAACTATGGAATGACAGACAGCTCAAACTACTATGTAA 420
Db 361 GGTCTACATGCTAATGGGAACTATGGAATGACAGACAGCTCAAACTACTATGTAA 420
QY 421 AGACCCATATTTAAGTTTACCTACAGTACCTGCTGCTGGAATGCCCTCAGGCAC 480
Db 421 AGACCCATATTTAAGTTTACCTACAGTACCTGCTGCTGGAATGCCCTCAGGCAC 480
QY 481 AAACCAATATGACCCCTATTCACATCCAGACAGCCCAAGACCGGCAAGCTTGCTT 540
Db 481 AAACCAATATGACCCCTATTCACATCCAGACAGCCCAAGACCGGCAAGCTTGCTT 540
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTCTGACAGTATGAGAAAGCACTATAC 600
Db 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTCTGACAGTATGAGAAAGCACTATAC 600
QY 601 ACCAATATGATGAGCTCAAAAGTCTCAAAATCAGCAATATTTCCGCTTACATGGA 660
Db 601 ACCAATATGATGAGCTCAAAAGTCTCAAAATCAGCAATATTTCCGCTTACATGGA 660
QY 661 TAATTACCTCAAGAAAGTCAATCAAGTGAAGAAAGCAAGGCTATTAACCTACTCAC 720
Db 661 TAATTACCTCAAGAAAGTCAATCAAGTGAAGAAAGCAAGGCTATTAACCTACTCAC 720
QY 721 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
Db 721 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
QY 781 TTAACAATNAGAGCAATAGTGGCTATCCAGACAGATGAATGCAAGTGGCTTACCAAT 840
Db 781 TTAACAATNAGAGCAATAGTGGCTATCCAGACAGATGAATGCAAGTGGCTTACCAAT 840
QY 841 TGTGATGTTTAAAGTAAAGTCAATGCTCCAGTAGAGACAGCCATCAAGTAA 900
Db 841 TGTGATGTTTAAAGTAAAGTCAATGCTCCAGTAGAGACAGCCATCAAGTAA 900
QY 901 TGTTCCTTCGGAATTAACCAAGTAGAAACAAACCGGAGCTGGGATCAACTATGAA 960
Db 901 TGTTCCTTCGGAATTAACCAAGTAGAAACAAACCGGAGCTGGGATCAACTATGAA 960
QY 961 ACCGATCAAGACTATGCTCCCTGCTGAGTAGCGTGTCTACGATCAACTGCTACTAT 1020
Db 961 ACCGATCAAGACTATGCTCCCTGCTGAGTAGCGTGTCTACGATCAACTGCTACTAT 1020
QY 1021 CGTTACAGTAGAGCCCTATTAACCTGGAGCAATACTCTGTTTAAAGTGGATAG 1080
Db 1021 CGTTACAGTAGAGCCCTATTAACCTGGAGCAATACTCTGTTTAAAGTGGATAG 1080
QY 1081 GGGCTACTTTGGCAACATCACTTGCATATAGCCCTGCAACAATGCGGAAGCTCCAC 1140

```

```

Db 1081 GGGCTACTTTGGCAACATCACTTGGCAATAGCCCTGCAACAATGCGGAAGCTCCAC 1140
QY 1141 CGTGAACCTTAACCAAGTGGAGCTCAACCGCGCAAGACTTTCTAAATGCTAG 1200
Db 1141 CGTGAACCTTAACCAAGTGGAGCTCAACCGCGCAAGACTTTCTAAATGCTAG 1200
QY 1201 AATGCACTACCAAGTATTCAGTACTCAAAATGCCATTTCAAGTAAACCAAGCAATCAGA 1260
Db 1201 AATGCACTACCAAGTATTCAGTACTCAAAATGCCATTTCAAGTAAACCAAGCAATCAGA 1260
QY 1261 CAAAAAATATGAGCAAGTAGTGAAGATGGCTGCTTACGCTGCTTTGCAAAATG 1320
Db 1261 CAAAAAATATGAGCAAGTAGTGAAGATGGCTGCTTACGCTGCTTTGCAAAATG 1320
QY 1321 TGAAGCTTCTAATAACCAATGATATCATTAAGTCAAGCTTATGATGGAGTGA 1380
Db 1321 TGAAGCTTCTAATAACCAATGATATCATTAAGTCAAGCTTATGATGGAGTGA 1380
QY 1381 AGAGTTCTTAATGTCGGAAGTGTGCAATGGAAGGAAGACAGCCTATATGATGACGA 1440
Db 1381 AGAGTTCTTAATGTCGGAAGTGTGCAATGGAAGGAAGACAGCCTATATGATGACGA 1440
QY 1441 CATGATGAAGAAAGCTTTGATCTTATGAACTGGACGAATGCCCTATCTTGGCTCCC 1500
Db 1441 CATGATGAAGAAAGCTTTGATCTTATGAACTGGACGAATGCCCTATCTTGGCTCCC 1500
QY 1501 TCAGGCTGTAAAGCAAGCAAGCTTACTATATACAGAGGGAATGAAGAACCATCA 1560
Db 1501 TCAGGCTGTAAAGCAAGCAAGCTTACTATATACAGAGGGAATGAAGAACCATCA 1560
QY 1561 GACCTCTCAATTTGTAGCACTGATGAACTATTTGCTGCTATACCGTAATATTTCAAT 1620
Db 1561 GACCTCTCAATTTGTAGCACTGATGAACTATTTGCTGCTATACCGTAATATTTCAAT 1620
QY 1621 GGCTGATGAGCAAGCTATTTCAACCGTCTGACCACTTGTAGGAATGGCTTACGGT 1680
Db 1621 GGCTGATGAGCAAGCTATTTCAACCGTCTGACCACTTGTAGGAATGGCTTACGGT 1680
QY 1681 CGCTGCCAAGTTTACCGCTATGATGACCTGCTGGAAGGAAGCAATCCAGAA 1740
Db 1681 CGCTGCCAAGTTTACCGCTATGATGACCTGCTGGAAGGAAGCAATCCAGAA 1740
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATGTAATTTAAAAATGCTCTG 1800
Db 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATGTAATTTAAAAATGCTCTG 1800
QY 1801 TTCTACGTGGAAGCTCACTGCTCTACACACACACCCCATCAACTGAAAGTCAAGCTCAT 1860
Db 1801 TTCTACGTGGAAGCTCACTGCTCTACACACACACCCCATCAACTGAAAGTCAAGCTCAT 1860
QY 1861 ATCAGATATTCACCTCAAGCTGATGCTCAACCACTCCAGACAAATATATGATGAC 1920
Db 1861 ATCAGATATTCACCTCAAGCTGATGCTCAACCACTCCAGACAAATATATGATGAC 1920
QY 1921 TACCAATCTTAACCAATATAGCAACATCAATATACACCCCTGATCAACAAATCAGAA 1980
Db 1921 TACCAATCTTAACCAATATAGCAACATCAATATACACCCCTGATCAACAAATCAGAA 1980
QY 1981 TCCTCAACGACACACCA 1999
Db 1981 TCCTCAACGACACACCA 1999

```

RESULT 2
 us-09-765-271-1
 ; Sequence 1, Application us/09765271
 ; GENERAL INFORMATION:
 ; APPLICANT: Choi et. al.
 ; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
 ; NUMBER OF SEQUENCES: 452
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue


```

? CITY: Rockville
? STATE: Maryland
? COUNTRY: USA
? ZIP: 20850
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage
? COMPUTER: HP Vectra 486/33
? OPERATING SYSTEM: MSDOS version 6.2
? SOFTWARE: ASCII Text
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/765,271
? FILING DATE: 22-Jan-2001
? CLASSIFICATION: <Unknown>
?
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 09/536,784
? FILING DATE: <Unknown>
? APPLICATION NUMBER: 08/961,083
? FILING DATE: OCT-30-1997
?
? ATTORNEY/AGENT INFORMATION:
? NAME: Michelle S. Marks
? REGISTRATION NUMBER: 41,971
? REFERENCE/DOCKET NUMBER: PB340P3
?
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (301) 309-8504
? TELEFAX: (301) 309-8512
?
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1999 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
?
? SEQUENCE DESCRIPTION: SEQ ID NO: 1:
? US-09-765-271-1

```

```

Query Match      100.0%; Score 1999; DB 30; Length 1999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TAAATCTAGACATAAATAATCACTGATGCTGATGGGTTCTGAAGCCGCTCAA 60
DB 1 TAAATCTAGACATAAATAATCACTGATGCTGATGGGTTCTGAAGCCGCTCAA 60
QY 61 TGCCCAAGCTAATGATATTCGCCACAGATTTGGTTAAGGCAATCGTTCTATGAAGCA 120
DB 61 TGCCCAAGCTAATGATATTCGCCACAGATTTGGTTAAGGCAATCGTTCTATGAAGCA 120
QY 121 TCGCTTTCTGGACACAGGGGGGATGATACCATCGGATCTGGGACTTCTTGGCGAA 180
DB 121 TCGCTTTCTGGACACAGGGGGGATGATACCATCGGATCTGGGACTTCTTGGCGAA 180
QY 181 TCTGCAAGCAATTCCTCCAAAGGTGATCACTCTCACCACCAAGTGTGATTAAGTGCAC 240
DB 181 TCTGCAAGCAATTCCTCCAAAGGTGATCACTCTCACCACCAAGTGTGATTAAGTGCAC 240
QY 241 TTAATCTTCACTTGCAGCTTCGACAGCACTATTTCTGTAAGGCTCAGGAAGCTTGTT 300
DB 241 TTAATCTTCACTTGCAGCTTCGACAGCACTATTTCTGTAAGGCTCAGGAAGCTTGTT 300
QY 301 AGCGATTCACTTGAACAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360
DB 301 AGCGATTCACTTGAACAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360
QY 361 GGTCTACATGCTCTAATGGGAATGATGGAATGACAGACAGCTCAAAATCTATGTTAA 420
DB 361 GGTCTACATGCTCTAATGGGAATGATGGAATGACAGACAGCTCAAAATCTATGTTAA 420
QY 421 AGACCTCAATTAATTAAGTTACCTCAGTTAGCCTTGCTGGCTGGAATGCTTCAGGACC 480
DB 421 AGACCTCAATTAATTAAGTTACCTCAGTTAGCCTTGCTGGCTGGAATGCTTCAGGACC 480
QY 481 AAACCAATATGACCCCTATTACATCCAGAAAGCCGCAAGCCGCAAGCAAGCTTGCTTT 540
DB 481 AAACCAATATGACCCCTATTACATCCAGAAAGCCGCAAGCCGCAAGCAAGCTTGCTTT 540

```

```

DB 481 AAACCAATATGACCCCTATTACATCCAGAAAGCCGCAAGCCGCAAGCAAGCTTGCTTT 540
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTGCTGAACAGATGAGAAAGACGATCAATAC 600
DB 541 ATCTGAATGAAAAATCAAGGCTACATCTGCTGAACAGATGAGAAAGACGATCAATAC 600
QY 601 ACCAATTACTGATGAGTACCAAAAGCTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGA 660
DB 601 ACCAATTACTGATGAGTACCAAAAGCTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGA 660
QY 661 TAATTACTCAGGAATCATCATCAAGTTGAGAAAGAAAGAAAGCAAGCTATACCTACAC 720
DB 661 TAATTACTCAGGAATCATCATCAAGTTGAGAAAGAAAGAAAGCAAGCTATACCTACAC 720
QY 721 AACTGGATGATGCTACACAAATGTAAGCAAGAAAGCTCAAAATATCTGCTGGATAT 780
DB 721 AACTGGATGATGCTACACAAATGTAAGCAAGAAAGCTCAAAATATCTGCTGGATAT 780
QY 781 TTACAATACAGACGAATACGTTGCCATCCAGAGATGAATTCGAAGTCTGCTTACCAT 840
DB 781 TTACAATACAGACGAATACGTTGCCATCCAGAGATGAATTCGAAGTCTGCTTACCAT 840
QY 841 TGTGATGTTTCTAACGCTAAGATCATTTGCCAGCTAGAGACAGCCATCATGTAAGTAA 900
DB 841 TGTGATGTTTCTAACGCTAAGATCATTTGCCAGCTAGAGACAGCCATCATGTAAGTAA 900
QY 901 TGTTCCTCGGAATTAACCAAGCAGTAAAGCAAAACCGGAGCTGGGATCAACTATGAA 960
DB 901 TGTTCCTCGGAATTAACCAAGCAGTAAAGCAAAACCGGAGCTGGGATCAACTATGAA 960
QY 961 ACCGATCAGACATATGCTCTGCTGGCTGGAGTAGCGGTCTACAGATTCAGCTCTACTAT 1020
DB 961 ACCGATCAGACATATGCTCTGCTGGCTGGAGTAGCGGTCTACAGATTCAGCTCTACTAT 1020
QY 1021 CGTTCAGATGAGCCCTATTAATACCTCGGAGCAAAATCTCTTATTAATCTGGGATAG 1080
DB 1021 CGTTCAGATGAGCCCTATTAATACCTCGGAGCAAAATCTCTTATTAATCTGGGATAG 1080
QY 1081 GGGCTACTTTGGCAACATCACTCTGCAATAGCGCCCTGCAACAAATCGGAAACGTCGAGC 1140
DB 1081 GGGCTACTTTGGCAACATCACTCTGCAATAGCGCCCTGCAACAAATCGGAAACGTCGAGC 1140
QY 1141 CGTGGAACTCTAATCAAGAGTGGAGCTCAACCGGCAAGCTTCTTAATAGTCTTAGG 1200
DB 1141 CGTGGAACTCTAATCAAGAGTGGAGCTCAACCGGCAAGCTTCTTAATAGTCTTAGG 1200
QY 1201 AATGACATACCAAGTATTAATCACTCAATAGCCATTTCAAGTAACCAACCGAATAGA 1260
DB 1201 AATGACATACCAAGTATTAATCACTCAATAGCCATTTCAAGTAACCAACCGAATAGA 1260
QY 1261 CAAAAAATATGAGCAAGTATGAAAGATGGCTGCTAGCTGCTGCTTGGCAATGG 1320
DB 1261 CAAAAAATATGAGCAAGTATGAAAGATGGCTGCTAGCTGCTGCTTGGCAATGG 1320
QY 1321 TGAAGCTTACTATAACCAATGTATATCCATTAAGTCTGCTTAGTGAAGGAGTGAATA 1380
DB 1321 TGAAGCTTACTATAACCAATGTATATCCATTAAGTCTGCTTAGTGAAGGAGTGAATA 1380
QY 1381 AGAGTCTCTAATGCGGAAGCTGCTGCAATGAAGAAAGCAAGCAAGCAAGCAAGCAAGCA 1440
DB 1381 AGAGTCTCTAATGCGGAAGCTGCTGCAATGAAGAAAGCAAGCAAGCAAGCAAGCAAGCA 1440
QY 1441 CATGATGAAAAACAGTCTGACTTATGGAAGCTGAGCAAGAAATGCTATGTTGGCTGCC 1500
DB 1441 CATGATGAAAAACAGTCTGACTTATGGAAGCTGAGCAAGAAATGCTATGTTGGCTGCC 1500
QY 1501 TCAGGCTGTGTAAGCAAGAACTCTAATATACAGAGAGAGAAATGAAAAACCATCAA 1560
DB 1501 TCAGGCTGTGTAAGCAAGAACTCTAATATACAGAGAGAGAAATGAAAAACCATCAA 1560
QY 1561 GACCTCTCAATTTGTACACCTGATGAAGCTATTTGCTGGCTATAGGCTTAATATTTCAAT 1620
DB 1561 GACCTCTCAATTTGTACACCTGATGAAGCTATTTGCTGGCTATAGGCTTAATATTTCAAT 1620

```


QY	1621	GGCTGTATGGACAGGGCTATCTTAACGCTGACACACTTGAAGCCAMGGCTTAACGCT	1680
Dp	1621	GCGCTGTATGGACAGGGCTATCTTAACGCTGACACACTTGAAGCCAMGGCTTAACGCT	1680
QY	1681	CGCTGCCAAGATTACCGCTCTATGATGACCTACCTGTGTGAAAGGACCAATCCGAGAAGA	1740
Dp	1681	CGCTGCCAAGATTACCGCTCTATGATGACCTACCTGTGTGAAAGGACCAATCCGAGAAGA	1740
QY	1741	TTGGAAATATACCAAGAGGGGCTCTACAGAAATGAGAGATTGCTATTTTAAAAATGGTGCTCG	1800
Dp	1741	TTGGAAATATACCAAGAGGGGCTCTACAGAAATGAGAGATTGCTATTTTAAAAATGGTGCTCG	1800
QY	1801	TTTACAGCTGGAACTCAGCTCTCTCCACAACACCCCATTAATGSAAGTCAAGCTCATC	1860
Dp	1801	TTTACAGCTGGAACTCAGCTCTCTCCACAACACCCCATTAATGSAAGTCAAGCTCATC	1860
QY	1861	ATCAGATAGTTCACACTTCACAGTCTAGCTCAACCACTCCACACCAAAATAATAGTAGAC	1920
Dp	1861	ATCAGATAGTTCACACTTCACAGTCTAGCTCAACCACTCCACACCAAAATAATAGTAGAC	1920
QY	1921	TACCAATCTCTAACAAATATATACGACAAATCAATACACCCCTGTATCAACAAAATCAGAA	1980
Dp	1921	TACCAATCTCTAACAAATATATACGACAAATCAATACACCCCTGTATCAACAAAATCAGAA	1980
QY	1981	TTCTTCACACGACGACAAACGA 1999	
Dp	1981	TTCTTCACACGACGACAAACGA 1999	

RESULT 3
US-09-765-272-1

; Sequence 1, Application US/09765272

GENERAL INFORMATION:

APPLICANT: CHOI ET. AL.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Ke
CTRY: Poland

STATE: Maryland

COUNTRY: U

; ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33

COMPIER: HP VECIId 480/33
OPERATING SYSTEM: MSDOS ver

```
SOFTWARE: ASCII Text
```

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/765,272
FILING DATE: 03 JUL 2001

FILED DATE: 22-Jan-2001
CLASSIFICATION: ~~CONFIDENTIAL~~

CLASSIFICATION: CONTINUOUS

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/961,083

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Brockus & Anderson

NAME: BROOKES, A. ANDERS
REGISTRATION NUMBER: 36.373

REFERENCE/DOCKET NUMBER: PB340P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8
INFORMATION FOR SEO ID NO: 1.

```

; INITIALIZATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

```

LENGTH: 1999 base pairs

TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: 1 in ear

TOPOLOGI: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-765-272-1

Query Match	100.0%;	Score 1999;	DB 30;	Length 1999;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1999; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

OY	1	TAAATGTACACAACTAAAATATCAAGCATGTGCTAGCTTGGGTCTGAACGCCGGCTAA	60
Db	1	TAAATGTACGACAACTAAAATCACTCAATGTCTGACTTGGTCTGAAAGCCGGCTAA	60
OY	61	TGCCAAGCTAATATATATTTCCACAGATTTTGTTAAGGCAATCGTTTCTATGGAAGACA	120
Db	61	TGCCAAGCTAATATATTTCCACAGATTTTGTTAAGGCAATCGTTTCTATGGAAGACA	120
OY	121	TGCGTTCTTGACACACAGGGGGATTTGATACATCCCTATACCTGGGAGCTTTCTTGCGAA	180
Db	121	TGCGTTCTTGACACACAGGGGGATTTGATACATCCCTATACCTGGGAGCTTTCTTGCGAA	180
OY	181	TCTGCAAGAAGAAATTCCTCCAGGTGGATCACTCCACCCCAACAGTTGATTAAGTTGAC	240
Db	181	TCTGCAAGAAGAAATTCCTCCAGGTGGATCACTCCACCCCAACAGTTGATTAAGTTGAC	240
OY	241	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT	300
Db	241	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT	300
OY	301	AGGCATTTCAGTTTACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATTAATPA	360
Db	301	AGGCATTTCAGTTTACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATTAATPA	360
OY	361	GGTGTACATGTCTAATGGGAAGTATGGAAATGCAGACAGAGCTCAAAACTACTATATGTA	420
Db	361	GGTGTACATGTCTAATGGGAAGTATGGAAATGCAGACAGAGCTCAAAACTACTATATGTA	420
OY	421	AGACCTCAATATATTTAAGTTTACCTCAGTTAGCCTTGTGGCTGGAGATGCTCAGGACCC	480
Db	421	AGACCTCAATATATTTAAGTTTACCTCAGTTAGCCTTGTGGCTGGAGATGCTCAGGACCC	480
OY	481	AAACCAATATGACCCCTAATTCACATCCAGAGCAGCCCAAGACCGCCGAAACTTGGTCTT	540
Db	481	AAACCAATATGACCCCTAATTCACATCCAGAGCAGCCCAAGACCGCCGAAACTTGGTCTT	540
OY	541	ATCTGAAATGAAAAATCAGAGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	600
Db	541	ATCTGAAATGAAAAATCAGAGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	600
OY	601	ACCAATTAAGTATGAGACTACAAAGTCTCAATACAGAACTAATTAACCTGTTACATGGA	660
Db	601	ACCAATTAAGTATGAGACTACAAAGTCTCAATACAGAACTAATTAACCTGTTACATGGA	660
OY	661	TAAATTACCTCAGGAAGTCTCATCATCAATCAAGTTGAAGAAAGCAAGCTATTAACCTACAC	720
Db	661	TAAATTACCTCAGGAAGTCTCATCATCAATCAAGTTGAAGAAAGCAAGCTATTAACCTACAC	720
OY	721	AACGTGGATGGATGTCTACACAATGTAGACACAGAGCTCAAAAAACATCTGTGGATAT	780
Db	721	AACGTGGATGGATGTCTACACAATGTAGACACAGAGCTCAAAAAACATCTGTGGATAT	780
OY	781	TTTCAATATACAGAGAAATAGCTTGGCTATCCAGACGATGAAATTTGGCAAGTCTTACATAT	840
Db	781	TTTCAATATACAGAGAAATAGCTTGGCTATCCAGACGATGAAATTTGGCAAGTCTTACATAT	840
OY	841	TGTTGATGTTTCTTAACGGTAAAGTTCATTTGCCACGCTAGAGAGCAGGCATCACTCAAGTAA	900
Db	841	TGTTGATGTTTCTTAACGGTAAAGTTCATTTGCCACGCTAGAGAGCAGGCATCACTCAAGTAA	900
OY	901	TGTTTCTTGCGGAATTAACCAACAGCATTAAGAAACACCGGACTGGGATCAACTATATA	960
Db	901	TGTTTCTTGCGGAATTAACCAACAGCATTAAGAAACACCGGACTGGGATCAACTATATA	960
OY	961	ACCCATACAGACATATGCTCTGCTGGAGTACGGGTGCTACGAATTCAACTGCTACTAT	1020
Db	961	ACCCATACAGACATATGCTCTGCTGGAGTACGGGTGCTACGAATTCAACTGCTACTAT	1020
OY	1021	CGTTTACAGATGACCTTATACTACCTTGGGACAATTACTCTGTTTAACTGGGATAG	1080


```

|||||
Db 1021 CGTTCAGATGACCCCTATACCTCGGACAAATACCTGTTATATACCTGGAGTAG 1080
QY 1081 GGGCTACTTGGCAACATCACTTGCATATGCGCCCTGCAACATCGGAAACGCCAGC 1140
Db 1081 GGGCTACTTGGCAACATCACTTGCATATGCGCCCTGCAACATCGGAAACGCCAGC 1140
QY 1141 CGTGAACACTTAACAGGTGCGACTCAACCGCGCAAGACTTCTTAATAGGTAGC 1200
Db 1141 CGTGAACACTTAACAGGTGCGACTCAACCGCGCAAGACTTCTTAATAGGTAGC 1200
QY 1201 AATGACTACCCAGTATTCATCTACTCAATGCCATTTCAGTACACAAACCGAATCAGA 1260
Db 1201 AATGACTACCCAGTATTCATCTACTCAATGCCATTTCAGTACACAAACCGAATCAGA 1260
QY 1261 CAAAAAATATGAGCAAGTGTGAAAAGATGGCTGCTTACGCTCCCTTTGCAATATGG 1320
Db 1261 CAAAAAATATGAGCAAGTGTGAAAAGATGGCTGCTTACGCTCCCTTTGCAATATGG 1320
QY 1321 TGGAACTTACTATTAACCAATGTATATCCATTAAGTCTTTAGTATGAGTGGAGTGA 1380
Db 1321 TGGAACTTACTATTAACCAATGTATATCCATTAAGTCTTTAGTATGAGTGGAGTGA 1380
QY 1381 AGAGTCTCTAATATGCGAACTCGTGCATGAAGAAACAGACCTTATATGATGACCGA 1440
Db 1381 AGAGTCTCTAATATGCGAACTCGTGCATGAAGAAACAGACCTTATATGATGACCGA 1440
QY 1441 CATGATGAAAAACAGTCTTGAATATGAGACGTGAGCAAAATGCCATTTGCTGGCTCC 1500
Db 1441 CATGATGAAAAACAGTCTTGAATATGAGACGTGAGCAAAATGCCATTTGCTGGCTCC 1500
QY 1501 TCAGGCTGTATTAACAGAACTCTTAATCTATACAGACGAGAAATTTGAAAAACCATCAA 1560
Db 1501 TCAGGCTGTATTAACAGAACTCTTAATCTATACAGACGAGAAATTTGAAAAACCATCAA 1560
QY 1561 GACCTCTCAATTTGTAGACACTGATGACTATTTGCTGCTATACGGTAAATTTTAAT 1620
Db 1561 GACCTCTCAATTTGTAGACACTGATGACTATTTGCTGCTATACGGTAAATTTTAAT 1620
QY 1621 GGGTGTATGAGACAGGCTATTTCTTAACCGTGTGACACCACTTGTAGGCAATGGCTTACGGT 1680
Db 1621 GGGTGTATGAGACAGGCTATTTCTTAACCGTGTGACACCACTTGTAGGCAATGGCTTACGGT 1680
QY 1681 CGGTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTGAAGAAACCAATCCAGAGA 1740
Db 1681 CGGTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTGAAGAAACCAATCCAGAGA 1740
QY 1741 TTGGAATATACAGAGGGGCTTACAGAAATGGAATTTGTAATTTAAATTTGCTGCTG 1800
Db 1741 TTGGAATATACAGAGGGGCTTACAGAAATGGAATTTGTAATTTAAATTTGCTGCTG 1800
QY 1801 TTCTAGCTGAGACCTCACTCTCTCACAACAAACCCCATCACTGAAGAAATTTCAAGCTCATC 1860
Db 1801 TTCTAGCTGAGACCTCACTCTCTCACAACAAACCCCATCACTGAAGAAATTTCAAGCTCATC 1860
QY 1861 ATCAGATAGTTCAACTCTCAGAGTCTAGCTCAACCACTCAAGCAACAATATATAGTACGAC 1920
Db 1861 ATCAGATAGTTCAACTCTCAGAGTCTAGCTCAACCACTCAAGCAACAATATATAGTACGAC 1920
QY 1921 TACCAATCTTAACAAATATATACGCAACATCAATATACACCCCTGATCACAATAATCAGAA 1980
Db 1921 TACCAATCTTAACAAATATATACGCAACATCAATATACACCCCTGATCACAATAATCAGAA 1980
QY 1981 TCCTCAACGACGACAAACCA 1999
Db 1981 TCCTCAACGACGACAAACCA 1999

```

```

RESULT 4
US-60-029-960-81/c
; Sequence 81, Application US/60029960
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch

```

```

; TITLE OF INVENTION: Streptococcus pneumoniae polynucleotides and sequences
; NUMBER OF SEQUENCES: 1649
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/029,960
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEO ID NO: 81:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10383 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-60-029-960-81

```

```

Query Match 100.0%; Score 1999; DB 41; Length 10383;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TAAATCTAGACATATAAATCACTATTTGCTGACTTGGGTTCTAGACGCCGCTCAA 60
Db 9763 TAAATCTAGACATATAAATCACTATTTGCTGACTTGGGTTCTAGACGCCGCTCAA 9704
QY 61 TGCCCAAGCTAATGATTTCCACAGATTTGGTTAAGCAATGCTTCTATCGAAGACA 120
Db 9703 TGCCCAAGCTAATGATTTCCACAGATTTGGTTAAGCAATGCTTCTATCGAAGACA 9644
QY 121 TCGCTTCTTGACACAGGGGGATTTGATACCATCCGATCCCTGGACCTTCTTGCGCAA 180
Db 9643 TCGCTTCTTGACACAGGGGGATTTGATACCATCCGATCCCTGGACCTTCTTGCGCAA 9584
QY 181 TCTGCAAGCAATTTCCCTCAAGGTGATCACTCTACCCACAGTTGATTAAGTTGAC 240
Db 9583 TCTGCAAGCAATTTCCCTCAAGGTGATCACTCTACCCACAGTTGATTAAGTTGAC 9524
QY 241 TTACTTTCAACTTCCGACTTCCGACGACGATTTTCTGTAAGGCTGAGAGCTTGTT 300
Db 9523 TTACTTTCAACTTCCGACTTCCGACGACGATTTTCTGTAAGGCTGAGAGCTTGTT 9464
QY 301 AGCGATTCAGTTGAGCAAAAAGCAACCAAGCAAGAAATCTTGACTACTATATAATTA 360
Db 9463 AGCGATTCAGTTGAGCAAAAAGCAACCAAGCAAGAAATCTTGACTACTATATAATTA 9404
QY 361 GGTCTACATGCTTAATGGAATGGAATGACAGACAGCTCAAAATCTACTATATGTTAA 420
Db 9403 GGTCTACATGCTTAATGGAATGGAATGACAGACAGCTCAAAATCTACTATATGTTAA 9344
QY 421 AGACCTCAATATTTAAGTTTACTCTGAGTTAGCTTCTGCTGGTGGATTCCTTACGAC 480
Db 9343 AGACCTCAATATTTAAGTTTACTCTGAGTTAGCTTCTGCTGGTGGATTCCTTACGAC 9284

```



```

OY 481 AACCAATATGACCCCTATTCATCCAGACAGCCCAAGCCGGAAGCTTGCTTT 540
    |||
DB 9283 AAACCAATATGACCCCTATTCATCCAGACAGCCCAAGCCGGAAGCTTGCTTT 9224
OY 541 ATCTGAAATGAAAAATCAAGGCTACTCTCTGTGACAGCTATGAGAAAAGCAGTCAATAC 600
    |||
DB 9223 ATCTGAAATGAAAAATCAAGGCTACTCTCTGTGACAGCTATGAGAAAAGCAGTCAATAC 9164
OY 601 ACCAATATGATGACATCAAGCTCAAAAGTCAAAATCAGCAAGTAAATTCCTGCTTACATGGA 660
    |||
DB 9163 ACCAATATGATGACATCAAGCTCAAAAGTCAAAATCAGCAAGTAAATTCCTGCTTACATGGA 9104
OY 661 TAATTACCTCAAGAGATCATCATCAAGTTAGAGAAAGAGAGAGCTATTAACCTACTCAC 720
    |||
DB 9103 TAATTACCTCAAGAGATCATCATCAAGTTAGAGAAAGAGAGAGCTATTAACCTACTCAC 9044
OY 721 AACTGGGATGAGATGCTCTACACAAATGTAGACAGAGAGCTCAAAAACATCTGTGGATAT 780
    |||
DB 9043 AACTGGGATGAGATGCTCTACACAAATGTAGACAGAGAGCTCAAAAACATCTGTGGATAT 8984
OY 781 TTACATATACAGAGATACGTCGCTATCCAGATGAGATGAGATGAGATGAGATGAGATGAGAT 840
    |||
DB 8983 TTACATATACAGAGATACGTCGCTATCCAGATGAGATGAGATGAGATGAGATGAGATGAGAT 8924
OY 841 TGTGATGTTTCTTAACGAGTAAAGTCAATGCGCAGTAGAGAGAGCAGTCAAGTCAAGTAA 900
    |||
DB 8923 TGTGATGTTTCTTAACGAGTAAAGTCAATGCGCAGTAGAGAGAGCAGTCAAGTCAAGTAA 8864
OY 901 TGTTCCTTCCGAAATTAACCAAGCAGTAGAACAACCGCAGCTGGGGATCAACTATGAA 960
    |||
DB 8863 TGTTCCTTCCGAAATTAACCAAGCAGTAGAACAACCGCAGCTGGGGATCAACTATGAA 8804
OY 961 ACCGATACAGAGATGAGTCCGCTGAGTACGCTGAGTACGCTGAGTACGCTGAGTACGCTAT 1020
    |||
DB 8803 ACCGATACAGAGATGAGTCCGCTGAGTACGCTGAGTACGCTGAGTACGCTGAGTACGCTAT 8744
OY 1021 CGTTCACGATGAGCCCTATTAACCTGAGCAAAATACCTCTGTTTAACTGGGATAG 1080
    |||
DB 8743 CGTTCACGATGAGCCCTATTAACCTGAGCAAAATACCTCTGTTTAACTGGGATAG 8684
OY 1081 GGGCTACTTGGCAACATCAGCTTGAATACGCCCTGCAACAATGCGGCAAAAGCTCCAC 1140
    |||
DB 8683 GGGCTACTTGGCAACATCAGCTTGAATACGCCCTGCAACAATGCGGCAAAAGCTCCAC 8624
OY 1141 CGTGGAACTTAACCAAGTGGAGTCAACCGCCCAAGATCTTTCCTTAATGCTTAG 1200
    |||
DB 8623 CGTGGAACTTAACCAAGTGGAGTCAACCGCCCAAGATCTTTCCTTAATGCTTAG 8564
OY 1201 AATGCACTACCCAGATTAATCACTCAATGCAATTCATCAAGTAAACCAACGAATCAGA 1260
    |||
DB 8563 AATGCACTACCCAGATTAATCACTCAATGCAATTCATCAAGTAAACCAACGAATCAGA 8504
OY 1261 CAAAAAATATGAGAGAGTGAAGAAAGATGCTGCTTACGCTGCTTACGCTTTCGCAATG 1320
    |||
DB 8503 CAAAAAATATGAGAGAGTGAAGAAAGATGCTGCTTACGCTGCTTTCGCAATG 8444
OY 1321 TGGAACTTACATTAACCAATGATATATCATTAAGTCCGCTTACGAGGAGGAGAAA 1380
    |||
DB 8443 TGGAACTTACATTAACCAATGATATATCATTAAGTCCGCTTACGAGGAGGAGAAA 8384
OY 1381 AGAGTTCTCTAATGTCGGAATCGTGCCATGAAGGAAAGAGAGCTTATATGACGGA 1440
    |||
DB 8383 AGAGTTCTCTAATGTCGGAATCGTGCCATGAAGGAAAGAGAGCTTATATGACGGA 8324
OY 1441 CATGATGAAAGAGTCTGACTTATGGAAGTGGAGAAATGCTTCTGCTTGGCTCC 1500
    |||
DB 8323 CATGATGAAAGAGTCTGACTTATGGAAGTGGAGAAATGCTTCTGCTTGGCTCC 8264
OY 1501 TCAGGCTGTAAAGAGAGAACTCTAATATACAGAGAGGAAATGGAACCAATCA 1560
    |||
DB 8263 TCAGGCTGTAAAGAGAGAACTCTAATATACAGAGAGGAAATGGAACCAATCA 8204
OY 1561 GACCTCTCAATTTGAGACCTGATGAACATATTTGCTGGCTATAGCGCTAATATTCAT 1620

```

```

DB 8203 GACCTCTCAATTTGAGACAGCTGATGACTATTGCTGCTATAGCGCTAAATATTCAT 8144
OY 1621 GGCTGATGAGACAGCTATTTCAACCGCTGTGACACCACTTTGAGCAATAGGCTTACG 1680
    |||
DB 8143 GGCTGATGAGACAGCTATTTCAACCGCTGTGACACCACTTTGAGCAATAGGCTTACG 8084
OY 1681 CGCTGCCAAGTTTACCGCTGTATGATGACCTGCTGTGAAAGAGCAATCCAGAGA 1740
    |||
DB 8083 CGCTGCCAAGTTTACCGCTGTATGATGACCTGCTGTGAAAGAGCAATCCAGAGA 8024
OY 1741 TTGGAATATACAGAGAGGCTCTACAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 1800
    |||
DB 8023 TTGGAATATACAGAGAGGCTCTACAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 7964
OY 1801 TTCTACGTGGAAGTCACTGCTGCTCCACAAACACCCCATCAATCAAGTCAAGTCAATC 1860
    |||
DB 7963 TTCTACGTGGAAGTCACTGCTGCTCCACAAACACCCCATCAATCAAGTCAAGTCAATC 7904
OY 1861 ATCAGATAGTCAACTTCAAGTCTGCTCAACCACTCCAGAGCAAAATATATGATGAG 1920
    |||
DB 7903 ATCAGATAGTCAACTTCAAGTCTGCTCAACCACTCCAGAGCAAAATATATGATGAG 7844
OY 1921 TACCAATCCTAATATATACGCAACATCAATCAATCAACCCCTGATCAACAAATCAGA 1980
    |||
DB 7843 TACCAATCCTAATATATACGCAACATCAATCAATCAACCCCTGATCAACAAATCAGA 7784
OY 1981 TCCTCACCAGCAACCA 1999
    |||
DB 7783 TCCTCACCAGCAACCA 7765

```

```

RESULT 5
US-08-961-527-145/c
; Sequence 145, Application US/08961527
;
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961.527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36, 373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 145:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10711 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-961-527-145

```


Query Match 100.0%; Score 1999; DB 13; Length 10711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAATCTAGCAATATTAATCACTATGCTGACTTGGGTTCTGACGCCGCCGCA 60
DB 9767 TAAATCTAGCAATATTAATCACTATGCTGACTTGGGTTCTGACGCCGCCGCA 9708

QY 61 TGGCAAGCTAATGATATTCACAGATTTGTTAAGGCAATGTTCTATGACGCA 120
DB 9707 TGGCAAGCTAATGATATTCACAGATTTGTTAAGGCAATGTTCTATGACGCA 9648

QY 121 TCGCTTCTTGACACAGGGGGGATGATACCATCCGATCCGAGACTTCTTGCGCA 180
DB 9647 TCGCTTCTTGACACAGGGGGGATGATACCATCCGATCCGAGACTTCTTGCGCA 9588

QY 181 TCTGCAAGCAATTCCTCCAGGTGATCACTTCACCCAACTGTTAATGTTGAC 240
DB 9587 TCTGCAAGCAATTCCTCCAGGTGATCACTTCACCCAACTGTTAATGTTGAC 9528

QY 241 TTACTTTTCACTTCGACCTCGACCACTATTTCTCGTAAGGCTCAGGAAGTTGTT 300
DB 9527 TTACTTTTCACTTCGACCTTCGACCACTATTTCTCGTAAGGCTCAGGAAGTTGTT 9468

QY 301 AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATA 360
DB 9467 AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATA 9408

QY 361 GGTCTACATGTCTAATGGGCACTATGATGACAGCGCTCAAACTACTATGTTAA 420
DB 9407 GGTCTACATGTCTAATGGGCACTATGATGACAGCGCTCAAACTACTATGTTAA 9348

QY 421 AGACGCAATATTTAAGTTTACTCAGTACGCTTGGTGGTGAATGCTCAGGAC 480
DB 9347 AGACGCAATATTTAAGTTTACTCAGTACGCTTGGTGGTGAATGCTCAGGAC 9288

QY 481 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGCCGCGAACTTGCTT 540
DB 9287 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGCCGCGAACTTGCTT 9228

QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAACACTCATAC 600
DB 9227 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAACACTCATAC 9168

QY 601 ACCAATATGATGACTCAAGGCTCAAAATGCAAGTAATTAACCTGCTTAACATGA 660
DB 9167 ACCAATATGATGACTCAAGGCTCAAAATGCAAGTAATTAACCTGCTTAACATGA 9108

QY 661 TAATTAACCTCAAGGATCATCAATCAAGTTGAAGAAAGAGGCTATTAACCTACTCAC 720
DB 9107 TAATTAACCTCAAGGATCATCAATCAAGTTGAAGAAAGAGGCTATTAACCTACTCAC 9048

QY 721 AACTGGATGATGATCTACCAAAATGATGACCAAGAGCTCAAAAACATCTGTTGGATAT 780
DB 9047 AACTGGATGATGATCTACCAAAATGATGACCAAGAGCTCAAAAACATCTGTTGGATAT 8988

QY 781 TTACATATACAGAGTAATGCTATCCAGAGATGAAATGCAAGTCCCTCTACCAT 840
DB 8987 TTACATATACAGAGTAATGCTATCCAGAGATGAAATGCAAGTCCCTCTCTACCAT 8328

QY 841 TGTGATGTTTCTAAGGTAAGATCAATGCTAGGCTAGAGACGCCATCACTCAAGTAA 900
DB 8927 TGTGATGTTTCTAAGGTAAGATCAATGCTAGGCTAGAGACGCCATCACTCAAGTAA 8868

QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGGATCAACTATGAA 960
DB 8867 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGGATCAACTATGAA 8808

QY 961 ACCGATCAGAGTATGCTCTGCTGAGTACGAGTGTCTAGATTAACCTGCTACTAT 1020
DB 8807 ACCGATCAGAGTATGCTCTGCTGAGTACGAGTGTCTAGATTAACCTGCTACTAT 8748

QY 1021 CGTTCAGATGAGCCCTATTAACCTCTGGGACAAATACCTCTGTTTATTAACCTGGATAG 1080
DB 8747 CGTTCAGATGAGCCCTATTAACCTCTGGGACAAATACCTCTGTTTATTAACCTGGATAG 8688

QY 1081 GGGCTACTTTGGCAACATCACTTGCATATGCGCTGCAACATCGGAAACGCTCCAGC 1140
DB 8687 GGGCTACTTTGGCAACATCACTTGCATATGCGCTGCAACATCGGAAACGCTCCAGC 8628

QY 1141 CGTGAACCTCTAACAAGGCTGGAGCTCAACCGCCGCAACACTTTCCTTAATGGTCTAGG 1200
DB 8627 CGTGAACCTCTAACAAGGCTGGAGCTCAACCGCCGCAACACTTTCCTTAATGGTCTAGG 8568

QY 1201 AATCGACTACCAAGTATTCATCTCACTCAATGCCATTTCAAGTAACAACCAAGTACGA 1260
DB 8567 AATCGACTACCAAGTATTCATCTCACTCAATGCCATTTCAAGTACCAACCAAGTACGA 8508

QY 1261 CAAAAATATGAGCAAGTATGTAAGAAATGCTGCTGCTTACGCTCCCTTTGCAATGG 1320
DB 8507 CAAAAATATGAGCAAGTATGTAAGAAATGCTGCTGCTTACGCTCCCTTTGCAATGG 8448

QY 1321 TGGAACTTACTATTAACCAATGATATATCATTAAGTCTGTTAGTGGAGTGAAA 1380
DB 8447 TGGAACTTACTATTAACCAATGATATATCATTAAGTCTGTTAGTGGAGTGAAA 8388

QY 1381 AGAGTTCTTAATGTGGAACCTGCTGCATGAAGAAACGACAGCCTATATGATGACGA 1440
DB 8387 AGAGTTCTTAATGTGGAACCTGCTGCATGAAGAAACGACAGCCTATATGATGACGA 8328

QY 1441 CATGATGAAAAACAGTCTTGACTTATGGAACCTGGAACGAAATGCTATCTTGCTGCC 1500
DB 8327 CATGATGAAAAACAGTCTTGACTTATGGAACCTGGAACGAAATGCTATCTTGCTGCC 8268

QY 1501 TCAAGCTGTAAAAAGAGACCTCTAATCTATACAGAGAGAAATGGAACCAATCA 1560
DB 8267 TCAAGCTGTAAAAAGAGACCTCTAATCTATACAGAGAGAAATGGAACCAATCA 8208

QY 1561 GACCTCTCAATTTGTAGCACCTGATGAATGATTTGCTGCTATACGCTAATATTCAT 1620
DB 8207 GACCTCTCAATTTGTAGCACCTGATGAATGATTTGCTGCTATACGCTAATATTCAT 8148

QY 1621 GGTGTATGACAGGCTATTTCTAACCCTGTGACACCACTTGTAGGCAATGGCTTACGGT 1680
DB 8147 GGTGTATGACAGGCTATTTCTAACCCTGTGACACCACTTGTAGGCAATGGCTTACGGT 8088

QY 1681 CGCTGCCAAGTTTACCGCTCTATGATGACCTGCTGGAAGAAACATCCAGAGA 1740
DB 8087 CGCTGCCAAGTTTACCGCTCTATGATGACCTGCTGGAAGAAACATCCAGAGA 8028

QY 1741 TTGGAATATACAGAGGGCTCTACAGAAATGAGAAATTCGTATTTAAAAATGGCTCG 1800
DB 8027 TTGGAATATACAGAGGGCTCTACAGAAATGAGAAATTCGTATTTAAAAATGGCTCG 7968

QY 1801 TTCTAGTGAACCTACCTGCTCCACAAACACCCCATCACTGAAAGTTCAAGCTCATC 1860
DB 7967 TTCTAGTGAACCTACCTGCTCCACAAACACCCCATCACTGAAAGTTCAAGCTCATC 7908

QY 1861 ATCAGATAGTTCAACTTCAAGTCTAGCTCAACACCTCCAGACAAATATATGATACAG 1920
DB 7907 ATCAGATAGTTCAACTTCAAGTCTAGCTCAACACCTCCAGACAAATATATGATACAG 7848

QY 1921 TACCAATCTTAACATATATACGAACATCAAAATACAAACCTTGATCAACAAATCGAA 1980
DB 7847 TACCAATCTTAACATATATACGAACATCAAAATACAAACCTTGATCAACAAATCGAA 7788

QY 1981 TCTCTCAACGACCAACCA 1999
DB 7787 TCTCTCAACGACCAACCA 7769

RESULT 6
US-09-583-110-1312
; Sequence 1312, Application US/09583110
; GENERAL INFORMATION:

Db 1959 ttctacgtggaactcacctgctccacaacaccccccaactcaactgaagttcaagctcatic 2018
Qy 1861 ATGAGATAGTTCACACTCTAGCTAGCTACACCACTCCAGCACAATATAGTAGTACGAC 1920
Db 2019 atcagatagttcaacttcaagctcagctcgaacacccacgaagaataatagtagaac 2078
Qy 1921 TACCAATCTTACAAATATATAGCAACAATCAATATCAACCCCTGATCAGACAAATCAGAA 1980
Db 2079 taccacatccataataatagcaacaataataacacccctgatacaacaataatcagaa 2138
Qy 1981 TCCTCAACCGACACACCA 1999
Db 2139 tcctcaacgagcaacaa 2157

RESULT 7

US-09-107-433-1102
Sequence 1102, Application US/09107433

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID

SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN

THE RAPIDITY

NUMBER OF SEQUENCES: 5206

CORRESPONDENCE ADDRESS:

ADDRESS: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: <Unknown>

OPERATING SYSTEM: <Unknown>

SOFTWARE: <Unknown>

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,433

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085131

FILING DATE: May 12, 1998

APPLICATION NUMBER: 60/051553

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-011

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 1102:

SEQUENCE CHARACTERISTICS:

LENGTH: 2166 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: misc.feature

LOCATION: (B) LOCATION 1..2166

SEQUENCE DESCRIPTION: SEQ ID NO: 1102:

US-09-107-433-1102

Query Match 82.1%; Score 1642; DB 15; Length 2166;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1992; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAAATCTAGACATAAATAATCACTCATCTGACTTGCGTTCTGACGCCGCTCAA 60
Db 165 TAAATCTAGACATAAATAATCACTCATCTGACTTGCGTTCTGACGCCGCTCAA 224
Qy 61 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATGCTTTCTATCGAAGCA 120
Db 225 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATGCTTTCTATCGAAGCA 284
Qy 121 TGCGTTCTTGACACAGGGGGGATTTGATACCATCCGATCTGAGAGCTTTCTTGCGCAA 180
Db 285 TGCGTTCTTGACACAGGGGGGATTTGATACCATCCGATCTGAGAGCTTTCTTGCGCAA 344
Qy 181 TGCGCAAGCAATTCCTCCAGAGTGATCACTCACCACCAAGTTGATTGATTGAC 240
Db 345 TGCGCAAGCAATTCCTCCAGAGTGATCACTCACCACCAAGTTGATTGATTGAC 404
Qy 241 TTACTTTTCAACTTCGACTTCGACGAGACTATTTTCGTGAAGCTTACGAGAGCTTGCT 300
Db 405 TTACTTTTCAACTTCGACTTCGACGAGACTATTTTCGTGAAGCTTACGAGAGCTTGCT 464
Qy 301 AGGATTCAGTTAGACAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 360
Db 465 AGGATTCAGTTAGACAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 524
Qy 361 GGTCTACATGCTAATGGAAGCTATGGAATGCAGACAGAGCTTCAAACTATGTTAA 420
Db 525 GGTCTACATGCTAATGGAAGCTATGGAATGCAGACAGAGCTTCAAACTATGTTAA 584
Qy 421 AGACCTCAATATTTAAGTTTACCTGATAGCTTCTGCTGCTGGAATGCTTACGAGACC 480
Db 585 AGACCTCAATATTTAAGTTTACCTGATAGCTTCTGCTGCTGGAATGCTTACGAGACC 644
Qy 481 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGAGCCGCAAGCTTGCTT 540
Db 645 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGAGCCGCAAGCTTGCTT 704
Qy 541 ATCTGAATGAAAAATCAAGGCTACATCTGCTGAAGCAAGTATGAGAAAGCAAGTCAATAC 600
Db 705 ATCTGAATGAAAAATCAAGGCTACATCTGCTGAAGCAAGTATGAGAAAGCAAGTCAATAC 764
Qy 601 ACCAATTAAGTATGAGTACAAAGTCAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 660
Db 765 ACCAATTAAGTATGAGTACAAAGTCAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 824
Qy 661 TAAATTAAGTATGAGTACAAAGTCAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 720
Db 825 TAAATTAAGTATGAGTACAAAGTCAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 884
Qy 721 AACTGGATGATGCTTACACAAATGTAGACCAAGAAAGTCAAAACATCTGTGGATAT 780
Db 885 AACTGGATGATGCTTACACAAATGTAGACCAAGAAAGTCAAAACATCTGTGGATAT 944
Qy 781 TTTCATATACAGAGAAATAGTGGCTATCCAGAGAGTGAATTCGAATTCGCTTCTACAT 840
Db 945 TTTCATATACAGAGAAATAGTGGCTATCCAGAGAGTGAATTCGAATTCGCTTCTACAT 1004
Qy 841 TGTGATGTTTCTAACGGTAAAGTCAATTCGCCAGCTTGAAGAGCAGCCATCAGTCAAGTAA 900
Db 1005 TGTGATGTTTCTAACGGTAAAGTCAATTCGCCAGCTTGAAGAGCAGCCATCAGTCAAGTAA 1064
Qy 901 TGTTCCTTGGAAATTAACCAAGAGTGAAGAAACCGGAGCTGGGATCACTATGAA 960
Db 1065 TGTTCCTTGGAAATTAACCAAGAGTGAAGAAACCGGAGCTGGGATCACTATGAA 1124
Qy 961 ACCGATACAGACTATGCTCTGCTGGAGTACAGGTCTAGAGTTCACCTGCTATAT 1020
Db 1125 ACCGATACAGACTATGCTCTGCTGGAGTACAGGTCTAGAGTTCACCTGCTATAT 1184
Qy 1021 CGTTACAGATGAGCCCTATATACCTGGGACAAATATCCCTGTTTATATACGTGGATAG 1080
Db 1185 CGTTACAGATGAGCCCTATATACCTGGGACAAATATCCCTGTTTATATACGTGGATAG 1244
Qy 1081 GGCGTACTTTGGCAACATCACTTGAATAGCGCTGCAACATTCGGAAACGTCCAGC 1140


```
|||||
Db 1245 GGGGACTTGGCAACATCATCCCTGCAATACGCCCTGCAACATGCGCAAGCTCCAC 1304
Qy 1141 CGTGGAACTCTAAACAGCTGGAGCTCAACGGCGCAAGACTTTCTTAAAGTCTAG 1200
Db 1305 CGTGGAACTCTAAACAGCTGGAGCTCAACGGCGCAAGACTTTCTTAAAGTCTAG 1364
Qy 1201 AATGCACTACCAAGATTTCACTACTCAATGCCATTTCAAGTAACACACGAATCAGA 1260
Db 1365 AATGCACTACCAAGATTTCACTACTCAATGCCATTTCAAGTAACACACGAATCAGA 1424
Qy 1261 CAAAAAATATGAGCAAGTAGTGAAGAGATGGCTGCTTACGCTGCCCTTGCAGATG 1320
Db 1425 CAAAAAATATGAGCAAGTAGTGAAGAGATGGCTGCTTACGCTGCCCTTGCAGATG 1484
Qy 1321 TGGACTTCTATTAACCAATATATATCATATAAGTCCGCTTATAGTAGAGAGTGA 1380
Db 1485 TGGAACTTATCTTAAACCAATATATATCATATAAGTCCGCTTATAGTAGAGAGTGA 1544
Qy 1381 AGAGTTCTCTATATGCGAACTGCGCATGGAAGAAAGACAGCCATATATGATGACGA 1440
Db 1545 AGAGTTCTCTATATGCGAACTGCGCATGGAAGAAAGACAGCCATATATGATGACGA 1604
Qy 1441 CATGATGAAAACAGTCTTGACTTATGGAAGTGGAGAAATGCCATATCTTGCTGCTCC 1500
Db 1605 CATGATGAAAACAGTCTTGACTTATGGAAGTGGAGAAATGCCATATCTTGCTGCTCC 1664
Qy 1501 TTAGGCTGGTAAACAGAGACCTCTACATATACAGACAGGAAATTTGAACACATCAA 1560
Db 1665 TTAGGCTGGTAAACAGAGACCTCTACATATACAGACAGGAAATTTGAACACATCAA 1724
Qy 1561 GACCTCTCAATTTTATGACACCTGATGAATATTTGCTGGCTATAGCGTAATATTCAT 1620
Db 1725 GACCTCTCAATTTTATGACACCTGATGAATATTTGCTGGCTATAGCGTAATATTCAT 1784
Qy 1621 GGCTGATGAGACAGCTATTTCAACCGTCTGACACCACTTGTAGCAATGGCCTTACG 1680
Db 1785 GGCTGATGAGACAGCTATTTCAACCGTCTGACACCACTTGTAGCAATGGCCTTACG 1844
Qy 1681 GGCTGCAAAAGTTTACCGCTGATGATGACCTGCTGCTGAGAGAAAGTCCAGAGA 1740
Db 1845 GGCTGCAAAAGTTTACCGCTGATGATGACCTGCTGCTGAGAGAAAGTCCAGAGA 1904
Qy 1741 TTGGAATATACAGAGGGCTCTACAGAAATGAGAAATTCGATTTAAATAATGAGCTCG 1800
Db 1905 TTGGAATATACAGAGGGCTCTACAGAAATGAGAAATTCGATTTAAATAATGAGCTCG 1964
Qy 1801 TTCTACGTGGAACCTGCTGCTGCTGACACACCCCATCACTGAAGTTCAAGCTCATC 1860
Db 1965 TTCTACGTGGAACCTGCTGCTGCTGACACACCCCATCACTGAAGTTCAAGCTCATC 2024
Qy 1861 ATCAATAGTAACTTACACTTACAGTCTGACACCACTCCAGACACAAATATATAGACG 1920
Db 2025 ATCAATAGTAACTTACACTTACAGTCTGACACCACTCCAGACACAAATATATAGACG 2084
Qy 1921 TACCAATCTTAATTAATAGCAATATCAATCAATCAATCCCTGATCAACAAATCAGA 1980
Db 2085 TACCAATCTTAATTAATAGCAATATCAATCAATCAATCCCTGATCAACAAATCAGA 2144
Qy 1981 TCCCTAACGACACACCA 1999
Db 2145 TCCCTAACGACACCA 2163

RESULT 8
PCT-US02-03987-9325
; SEQUENCE 9325, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; FILE REFERENCE: ELITRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
```

```
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FASTSEQ For Windows Version 4.0
; SEQ ID NO 9325
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2160)
; PCT-US02-03987-9325

Query Match 71.9%; Score 1438; DB 1; Length 2160;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 TAAATCTACGACAAATAAATCACTCATTTGCTGACTGGGTTCTGAACGGCGGTCAA 60
Db 159 taaatctacgacaataaatacaactcatgtctgacttgggtctgaaagccggtcaa 218
Qy 61 TGCCCAAGCTAATGATATTTCCACAGATTTGTTAAGCAATCGTTTATCGAAGACCA 120
Db 219 tgcccaagctaataatattccacagatttggtaaggaatcgtttctatcgaagacca 278
Qy 121 TCGCTTTCTGACACACAGGGGGATGATACATCCGTAATCTGAGAGCTTTCTGGCAA 180
Db 279 tcgctttctgacacacagggggaattgataccatccgtatccgtggagctttctggcaa 338
Qy 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTGACCCCAAGTATTAATTAATGATGAC 240
Db 339 tctgcaagcaatctccctcccaagggtgatacctgaccccaagtatttaatttaattgac 398
Qy 241 TTACTTTTCACTTGACCTTCCGACCACTATTTCTGTAAGCTCAGCAAGCTTGTT 300
Db 399 ttacttttcaacttgacttccgaccagactatttctgtaagctcaggaagcttggtt 458
Qy 301 AGCGATTCAGTTAGAAACAAAGCAACCAAGCAAGCAAGAAATCTTGACTACTATTAATA 360
Db 459 agcgatctcagttagaacaaagcaacccaagcaagaatcttgaactatataataaa 518
Qy 361 GGTCTACATGTTAATGGAAGACTATGGAATGCAGACAGCAAGCTCAATATGATGTA 420
Db 519 ggtctacatgtttaatgggaagactatgggaatgcagacagcaagctcaaatatgataa 578
Qy 421 AGACCTCAATTAATTAAGTTTACCTGATTAGCTTGTGCTGGAATGCTCAGGACAC 480
Db 579 agacctcaatatttaagtttacctgattagcttgtgctggaatgctcaggaagctcac 638
Qy 481 AAACCAATATGACCCCTATTTACATCTCAGACAGACGCCCAAGACCGCGAACTTGCT 540
Db 639 aaaccaatattgacccctatttcaatcagagaagcccaagaccgccaagcttgctt 698
Qy 541 ATCTGAAATGAATAATCAAGGCTACATCTGCTGAGAGTGTGAAGAGCAATCAATAC 600
Db 699 atctgaaatgaataatcaaggctacatctgctgagagtgatgaaagcaagcttaaac 758
Qy 601 ACCAATTAATGATGACTCAAGAGTCTCAATCAGCAAGCAATTAATTAATTAATTAAT 660
Db 759 accaatttaataatgactcaagagtctcaatcagcaagcaatataatcaccgtctaac 818
Qy 661 TAATTAATCTCAAGAAATGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 720
Db 819 taatttaactcaagaaatgataatcaatcaatcaatcaatcaatcaatcaatcaatcaat 878
Qy 721 AACTGGAGTATGATGCTTACCAATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 879 aactggagtatgattgcttaccatgataatgataatgataatgataatgataatgataat 938
Qy 781 TTACATACAGAGCAATACGTTGCTATCAGACGATGATGATGATGATGATGATGATGAT 840
Db 840 ttacatacagagcaatacgttgctatcagacgataatgataatgataatgataatgataat
```


Db 939 ttaacaatagacgaatagcttgccataccaagcagatgcaagtcgcttaccat 998
 QY 841 TGGTATGTTCTAAGCGTAAAGCATTTGCCGCTAGAGCGCCGCTAGTCAAGTAA 900
 Db 999 tctgaagttctcaacgaagtaacatctgcacgctagagcagcgacatcaagtaa 1058
 QY 901 TGTTCCTTCGGAATTAAACAGCAGTAGAAGAAACCGGAGCTGGGATCACTATGAA 960
 Db 1059 tgttctctcggaatttaaccaagcagtagaacaacacgagctgggagatcaactagaa 1118
 QY 961 ACCGATCAGAGACTATGCTCTGCTTGGAGTACGGTGTCTAGATTCAACTGCTACTAT 1020
 Db 1119 accgatcaacgactatgtctctgcttgagtagcgtgtctacgagtcacatcgccactat 1178
 QY 1021 CGTTACAGAGAGCCCTATTAAGTACCTGGGACAAATATGCTGTTATTAAGTGGAGTAG 1080
 Db 1179 cgttcaagaagagcccttaacacccctggacaataaccccgattataacatgaggtag 1238
 QY 1081 GGGCTACTTTGGACATCATCCTTGGCAATACGCCCTGCAAAATCGGAAACGTCGACG 1140
 Db 1239 gggctactttggcaacatcaacttgcaatagcgccttgcaacatcggaacgtccagc 1298
 QY 1141 CGTGGAACTCTAAGCAGTGGACTCAACCGGCCAAGACTTTCCTAATGCTTAGG 1200
 Db 1299 cgtgaaaccttaacaaagtcgactcaacgcgcgaagacttcttaaatgctcgcg 1358
 QY 1201 AATCGACTACCAAGTATTACTACTCAATGGCATTTCAAGTAAACCAACCGAATCAGA 1260
 Db 1359 aatcgactaccaagatatacactcaacatcgcatltaagaataacacacgaatcaga 1418
 QY 1261 CAAAAAATATGAGCAAGTAGTAAAAAGATGGCTGCTTACGCTGCCCTTTCGAAATGG 1320
 Db 1419 caaaaaatagcgcaagtagtgaagaagatgctgcttcaacgtccttgcaaatg 1478
 QY 1321 TGGAACTTACTATTAACCATGTATATCCATTAAGTGTCTTAGTATGGAGTGAATA 1380
 Db 1479 tggaaacttactataacccaatglatatccataaagtcgtctttagtgaagtagtaaa 1538
 QY 1381 AGAGTTCTCTAATGTCGGAAGTGTGTCGATGAAGAAACGACAGCCATATGATGACGA 1440
 Db 1539 agagttctctaaatgctggaactcgtgcatagaagaagaacacgacctatagtacgca 1598
 QY 1441 CATGATGAAAAAGTCTTGACTTATGAGACGAAATGCTATCTTGTGCTGCC 1500
 Db 1599 catgatagaacacgctctgtagttagtaagacgaaatgctctatcttgctgcacc 1658
 QY 1501 TCAGGCTGGTAAACGAGACCTCTAATCTATACAGAGAGAAATTCAAAACCAATCAA 1560
 Db 1659 tcaagctggtgtaaaacaggaaccttaactatacagacgagaaatltgaaacacacataa 1718
 QY 1561 GACCTCTCAATTTGTAGACCTGATGACTATTTGTGCTATACGCGTAAATATTCAAT 1620
 Db 1719 gaccttcaaatgttagacactatgaactatgtctggtataacgsgtaaatatcaat 1778
 QY 1621 GGGTATATGACAGGCTATTTCTAACCGTCTGACACCACTTGTAGCAANTGGCTTACGT 1680
 Db 1779 ggtgtatgagacgactatcttaacgctgacacacactgttagcgaatgaccttaccgt 1838
 QY 1681 CGGTGCCAAGTTTACCGCTCTATGATGACTCTGTCTGAAGAGAAATTCACAGAGA 1740
 Db 1839 cgttgcacaaagtttaccgctctatgactactcgttctgaaagaaagcaatccagagga 1898
 QY 1741 TTGCAATATACAGAGGGGCTTACAGAAATGAGAAATTCGTAATTTAAATAGTGTCTG 1800
 Db 1899 ttggaatataccagagggctctcaacgaaatgagaatctgtatcttaaaaaatggtctcg 1958
 QY 1801 TTCTACGTGGAACCTCACTGCTCTCACAAACCCCATCAACTGAAGTTCAAGCTCATC 1860
 Db 1959 ttctacgttgagagctcactcgtctcacacaaccccatcaactgaagttcaagctcacc 2018
 QY 1861 ATCAGATAGTTCAACTTACAGTCTAGCTACACCACTCCAGACAAATATATAGTACGAC 1920
 Db 2019 atcagatagttcaacttcaacagttcgttcaaccactccaagcacaaataatagtagac 2078

QY 1921 TACCAATCTTAACATATATAGCAACATCAATCAACCCCTGATGCACAAAATCAGAA 1980
 Db 2079 taccaatcttaacataataatctgcacaataccaataatacaacccctgatacaaaaaatcaga 2138
 QY 1981 TCCCTCAACGACACACACCA 1999
 Db 2139 tccctcaaccgacacaaacca 2157
 RESULT 9
 US-09-815-242-9325
 ; Sequence 9325, Application US/09815242
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9325
 ; LENGTH: 2160
 ; TYPE: DNA
 ; ORGANISM: Streptococcus pneumoniae
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(2160)
 ; US-09-815-242-9325
 Query Match 71.9%; Score 1438; DB 31; Length 2160;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTAGACAAATTAATCAATCACTGCTGACTGGTTCGAAAGCGCGCA 60
 Db 159 taaatctagacaataaataaactcaactcgtgacttgggttcgaaagcgcgca 218
 QY 61 TGCCCAAGCTAATGATTTCCACAGATTTGTTAAGGCAATGTTCTATTGCAAGACA 120
 Db 219 tgcaccaagctaatgatatctccacagattggttaaggaatcgtttctatcgaaagca 278
 QY 121 TGCTTCTTGACACAGGGGATTTGATACATCCGATCTCTGGAGCTTTCTTGCGCAA 180
 Db 279 tgcgttcttgacacaggggattgataccatcgtatccttggagcttcttgcgaa 338
 QY 181 TCTGCAAGCAATTCCTTCAAGGTGATCACTCTCACCACAGATTTGATTAAGTTGAC 240
 Db 339 tctgcaagaatctccctcaagtgatcatgctctcaactcaacagttgattgaagtac 398

Db 2079 taccaatcctaataataatgcacaatacaatcaaccctgatacaaaaatcagaa 2138
 QY 1981 TCCTCAACGACGACACCA 1999
 ||||||||||||||||
 Db 2139 tcccaacagcacacaacaa 2157

RESULT 11

US-60-061-998-596/C

Sequence 596, Application US/60061998

GENERAL INFORMATION:

APPLICANT: LAGACE, ROBERT E.

APPLICANT: CORLEY, NEIL C.

APPLICANT: RUSSO, FRANK D.

APPLICANT: HANN, AMY L.

APPLICANT: HEATH, JOE D.

APPLICANT: FINNEY, GREGORY L.

TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE

NUMBER OF SEQUENCES: 797

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/60/061,998

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: CERRONE, MICHAEL C.

REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PM-0006-2P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 596:

SEQUENCE CHARACTERISTICS:

LENGTH: 1033 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

IMMEDIATE SOURCE:

CLONE: SPN1C601

US-60-061-998-596

Query Match 71.9%; Score 1438; DB 45; Length 10333;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 198; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATTAATACTACTATGCTGACTTGGGTTCTGAACGCCGCTCAA 60
 ||||||||||||||||
 Db 9718 TAAATCTACGACATTAATACTACTATGCTGACTTGGGTTCTGAACGCCGCTCAA 9659
 QY 61 TGGCCAGGTATGATATTCACAGATTTGTTAGGCAATCGTTCTATCGAAGACCA 120
 ||||||||||||||||
 Db 9658 TGGCCAGGTATGATATTCACAGATTTGTTAGGCAATCGTTCTATCGAAGACCA 9599
 QY 121 TCGCTTCTTGACACAGGAGGATGATACCATCGTATCTGAGAGCTTTCTTGCGCAA 180
 ||||||||||||||||
 Db 9598 TCGCTTCTTGACACAGGAGGATGATACCATCGTATCTGAGAGCTTTCTTGCGCAA 9539
 QY 181 TCTGCAAGCAATTCCTCCAGAGGTGATCAACTCTACCCCAAGTTGATTAAGTTGAC 240
 ||||||||||||||||

Db 9538 TCTGCAAGCAATTCCTCCAGAGGTGATCAGCTCTACTCAACAGTTGATTAAGTTGAC 9479
 QY 241 TTAATCTTCAACTTGCATCTCCGACGACTATTTCTGTAAGCGTCAGAACCTTGTT 300
 ||||||||||||||||
 Db 9478 TTAATCTTCAACTTGCATCTCCGACGACTATTTCTGTAAGCGTCAGAACCTTGTT 9419
 QY 301 AGCGATTGAGTATGAAACAAAGCAACGACGAAGAATCTTGACTATTAATAA 360
 ||||||||||||||||
 Db 9418 AGCGATTGAGTATGAAACAAAGCAACGACGAAGAATCTTGACTATTAATAA 9359
 QY 361 GGTCTACATGCTTAATGGCACTATGGAATGACAGACAGCTCAAACTATGATG 420
 ||||||||||||||||
 Db 9358 GGTCTACATGCTTAATGGCACTATGGAATGACAGACAGCTCAAACTATGATG 9299
 QY 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGCGGTGAGATGCTAGGAC 480
 ||||||||||||||||
 Db 9298 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGCGGTGAGATGCTAGGAC 9239
 QY 481 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGACCGCAAACTTGCTTT 540
 ||||||||||||||||
 Db 9238 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGACCGCAAACTTGCTTT 9179
 QY 541 ATCTGAATGAAAAAATCAAGGCTACATCTCTGTAACAGTATGAAAGCACTAATAC 600
 ||||||||||||||||
 Db 9178 ATCTGAATGAAAAAATCAAGGCTACATCTCTGTAACAGTATGAAAGCACTAATAC 9119
 QY 601 ACCAATTACTGATGAGCTCAAAAGTCTCAAAATGACGAAGTATTCCTGCTACATGGA 660
 ||||||||||||||||
 Db 9118 ACCAATTACTGATGAGCTCAAAAGTCTCAAAATGACGAAGTATTCCTGCTACATGGA 9059
 QY 661 TAATTAACCTCAAGGAAGTATCAATCAAGTTGAAGAAGCAAGGCTAATACCTACTAC 720
 ||||||||||||||||
 Db 9058 TAATTAACCTCAAGGAAGTATCAATCAAGTTGAAGAAGCAAGGCTAATACCTACTAC 8999
 QY 721 AACTGGATGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAATCTGTGGATAT 780
 ||||||||||||||||
 Db 8998 AACTGGATGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAATCTGTGGATAT 8939
 QY 781 TTACATATACAGAGATATGCTGCTATCCAGAGATGATTCAGATGCTGCTACAT 840
 ||||||||||||||||
 Db 8938 TTACATATACAGAGATATGCTGCTATCCAGAGATGATTCAGATGCTGCTACAT 8879
 QY 841 TGTTCATGTTTCTAACGTAAGTCAATGCCCCAGCTAGAGAGCCGATCAGTCAAGTAA 900
 ||||||||||||||||
 Db 8878 TGTTCATGTTTCTAACGTAAGTCAATGCCCCAGCTAGAGAGCCGATCAGTCAAGTAA 8819
 QY 901 TGTTCCTTCGGAATTAACCAAGCACTAGAACAAACCGGACTGGGATCAACTATGAA 960
 ||||||||||||||||
 Db 8818 TGTTCCTTCGGAATTAACCAAGCACTAGAACAAACCGGACTGGGATCAACTATGAA 8759
 QY 961 ACCGATCACAGACTATGCTGCTGCTGAGTACGCTGATGATTCAGTCAACTGCTAT 1020
 ||||||||||||||||
 Db 8758 ACCGATCACAGACTATGCTGCTGCTGAGTACGCTGATGATTCAGTCAACTGCTAT 8699
 QY 1021 CGTTACGATGAGCCCTATTAACCTTGAGGACAAATCTCTGTTTAACCTGGGATAG 1080
 ||||||||||||||||
 Db 8698 CGTTACGATGAGCCCTATTAACCTTGAGGACAAATCTCTGTTTAACCTGGGATAG 8639
 QY 1081 GGGCTACTTTGGCAACATCACTTGCAATAGCCCTGCAACAATGCGGAAGCTCCAGC 1140
 ||||||||||||||||
 Db 8638 GGGCTACTTTGGCAACATCACTTGCAATAGCCCTGCAACAATGCGGAAGCTCCAGC 8579
 QY 1141 CGTGAAGACTTAACGAAGGTGGAGTCAACCGCGCAAGACTTCTTAATGCTAGG 1200
 ||||||||||||||||
 Db 8578 CGTGAAGACTTAACGAAGGTGGAGTCAACCGCGCAAGACTTCTTAATGCTAGG 8519
 QY 1201 AATGCACTACCAAGTATTCACACTCAATGCAATGCAATTCGAATCAACCAAGCAATCAGA 1260
 ||||||||||||||||
 Db 8518 AATGCACTACCAAGTATTCACACTCAATGCAATGCAATTCGAATCAACCAAGCAATCAGA 8459
 QY 1261 CAAAAAATATGAGCAAGTATGAAAAAGATGGCTGCTCTACGCTGCTTTGCAATAG 1320
 ||||||||||||||||
 Db 8458 CAAAAAATATGAGCAAGTATGAAAAAGATGGCTGCTCTACGCTGCTTTGCAATAG 8399

QY	1331	TGGAACCTACTAATAAACCAATGATATATCCATAAAGCGCTTACTGATATGGAGTGA	1380
Db	8338	TGGAACTTACTATTAACCAATGATATATCCATAAAGCGCTTACTGATATGGAGTGA	8339
QY	1381	AGAGCTCTCTAATATCTGGGAATCGTGCCATGGAAGAAACGACGCGTATATGATGACGA	1440
Db	8338	AGAGTCTCTAATATCTGGGAATCGTGCCATGGAAGAAACGACGCGTATATGATGACGA	8279
QY	1441	CATATATAAAACAGTCTTGACTATGAGCACTGGAGAAATGCGTATTCGCTTGGCTGCC	1500
Db	8278	CATATATAAAACAGTCTTGACTATGAGCACTGGAGAAATGCGTATTCGCTTGGCTGCC	8219
QY	1501	TCAGGCTGTATAAAACAGGAACCTCTAACATACAGACGAGAAATTTGAAACCAATCAA	1560
Db	8218	TCAGGCTGTATAAAACAGGAACCTCTAACATATACAGACGAGAAATTTGAAACCAATCAA	8159
QY	1561	GACCTCTCAATTTGTATGACACCTATATCACTATTTGTGTGCTATACGCTAAATATTCAT	1620
Db	8158	GACCTCTCAATTTGTATGACACCTATATCACTATTTGTGTGCTATACGCTAAATATTCAT	8099
QY	1621	GGCTGTATGACACAGGCAATCTAACCCTGTGACACACCTTGAGGCAATGGCGCTACGGT	1680
Db	8098	GGCTGTATGACACAGGCAATCTAACCCTGTGACACACCTTGAGGCAATGGCGCTACGGT	8039
QY	1681	CGCTGCCAAAGTTTACCGCTCTATGTATGACCTTACCTGTCTGAAGAGCAATCCAGAGA	1740
Db	8038	CGCTGCCAAAGTTTACCGCTCTATGTATGACCTTACCTGTCTGAAGAGCAATCCAGAGA	7979
QY	1741	TTTGAAATATACAGAGGGGCTCTACAGAAATGAGAAATCGATTTAAATAATGGGCTCG	1800
Db	7978	TTTGAAATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTAAATAATGGTCTCG	7919
QY	1801	TTTACGTGGAATCACTCTGCTCCACACACACCCCATCACTGAAGTCAAGCTATC	1860
Db	7918	TTTACGTGGAAGTCACTCTGCTCCACACACACCCCATCACTGAAGTCAAGCTATC	7859
QY	1861	ATCAGATAGTTCAACTTCAACAGTCTAGGTCACACACCTCCAAACCAATATATAGAGAC	1920
Db	7858	ATCAGATAGTTCAACTTCAACAGTCTAGGTCACACACCTCCAAACCAATATATAGAGAC	7799
QY	1921	TACCAATCTTAACATATATACGCAACAATCAAAATCAACCCCTGATTCACAAATTCAGAA	1980
Db	7798	TACCAATCTTAACATATATACGCAACAATCAAAATCAAAACCCCTGATTCACAAATTCAGAA	7739
QY	1981	TTCCTCAACGACGACCAACA 1999	
Db	7738	TTCCTCAACGACGACCAACA 7720	

RESULT 12
 US-08-116-541-3
 Sequence 3, Application US/08116541
 GENERAL INFORMATION:
 APPLICANT: Masure, H. Robert
 APPLICANT: Pearce, Barbara J.
 APPLICANT: Toumaen, Elaine
 TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
 TITLE OF INVENTION: CELLULAR VACCINES BASED THEREOF
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Klauber & Jackson
 STREET: 411 Hackensack Avenue
 CITY: Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

```

1 APPLICATION NUMBER: US/08/116,541
2 FILING DATE: 19930901
3 CLASSIFICATION: 424
4 ATTORNEY/AGENT INFORMATION:
5 NAME: Jackson Esq., David A.
6 REGISTRATION NUMBER: 26,742
7 REFERENCE/DOCKET NUMBER: 600-1-0699
8 TELECOMMUNICATION INFORMATION:
9 TELEPHONE: 201 487-8800
10 TELEFAX: 201 343-1684
11
12 TELEX: 133521
13
14 INFORMATION FOR SEQ ID NO: 3:
15 SEQUENCE CHARACTERISTICS:
16 LENGTH: 960 base pairs
17 TYPE: nucleic acid
18 STRANDEDNESS: both
19 TOPOLOGY: unknown
20 MOLECULE TYPE: DNA (genomic)
21 HYPOTHEetical: NO
22 ANTI-SENSE: NO
23 ORIGINAL SOURCE:
24 ORGANISM: Streptococcus pneumoniae
25 STRAIN: R6
26 IMMEDIATE SOURCE:
27 CLONE: SP8042
28 FEATURE:
29 NAME/KEY: CDS
30 LOCATION: 1..960
31
32 US-08-116-541-3

```

Query Match	39.8%	Score 796;	DB 5;	Length 960;
Best Local Similarity	99.7%	Pred. No. 0;		
Matches 946;	Conservative	0;	Mismatches 3;	Indels 0;
			Gaps	0

QY	1	TAAAATCAGACAAATAAAATCAACATCTGGACTTGGGTTCTAAAGCCGGCAA	60
Db	12	TAAAATCTAGACAAATAAAATCAACATCTGGACTTGGGTTCTAAAGCCGGCAA	71
QY	61	TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAAGCAATGCTTTATCGAAGACA	120
Db	72	TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAAGCAATGCTTTATCGAAGACA	131
QY	121	TTCGCTTCGTGCACACAGGGGGATGTGATACCATCCGATATCTGGAGACTTCTTGGCAA	180
Db	132	TTCGCTTCGTGCACACAGGGGGATGTGATACCATCCGATATCTGGAGACTTCTTGGCAA	191
QY	181	TCTGCACAAAGCAATTCCTCCACAGGTGATCAACTCTACCCCAACAGTTGATTAAGTTGAC	240
Db	192	TCTGCACAAAGCAATTCCTCCACAGGTGATCAACTCTACCAACAGTTGATTAAGTTGAC	251
QY	241	TTACTTTTCAACTTCGACTTCGCCACCGACTATTTCTTCGTAAAGGCTCAGAAAGCTTGTT	300
Db	252	TTACTTTTCAACTTCGACTTCGCCACCGACTATTTCTTCGTAAAGGCTCAGAAAGCTTGTT	311
QY	301	AGCATTTCAGTTAGTAACAAAAAGACCAACCAAGCAAGTAATCTTGACCTACATATTAATTA	360
Db	312	AGCATTTCAGTTAGTAACAAAAAGACCAACCAAGCAAGTAATCTTGACCTACATATTAATTA	371
QY	361	GGTCTACATGCTTAATGGGAATATGGAATGACAGACAGACCTCAAAACTACTATGCTAA	420
Db	372	GGTCTACATGCTTAATGGGAATATGGAATGACAGACAGCTCAAAACTACTATGCTAA	431
QY	421	AGACCTCAATTAATTAAGTTAACTCTAGTTAGACCTTCTGGCTGGGAATGCTCAGGACAC	480
Db	432	AGACCTCAATTAATTAAGTTAACTCTAGTTAGACCTTCTGGCTGGGAATGCTCAGGACAC	491
QY	481	AAACCAATATGACCCCTATTCATCTCAGAAAGCAGCCCAAGACCGCGAAACTTGCTTT	540
Db	492	AAACCAATATGACCCCTATTCATCTCAGAAAGCAGCCCAAGACCGCGAAACTTGCTTT	551
QY	541	ATCTGAAATGAAAAATCAGGCTACATCTCTGCTGAACAGATATGAGAAACAGTCAATAC	600


```

Db 552 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAAGATGAGAAAGAGTCATATAC 611
Qy 601 ACCAATTACTGATGAGTACCAAGTCTCAAAATCAGCAAGTAAATTAATCCCTGCTTACATGA 660
Db 612 ACCAATTACTGATGAGTACCAAGTCTCAAAATCAGCAAGTAAATTAATCCCTGCTTACATGA 671
Qy 661 TAATTAACGTCAGGAAGTATCATCAATCAAGTGAAGAAGAGGCTATTAACCTACTAC 720
Db 672 TAATTAACGTCAGGAAGTATCATCAATCAAGTGAAGAAGAGGCTATTAACCTACTAC 731
Qy 721 AACTGGAGATGATGCTTACACAATGTAGACCAAGAGCTCAAAAAATCTGTGGATAT 780
Db 732 AACTGGAGATGATGCTTACACAATGTAGACCAAGAGCTCAAAAAATCTGTGGATAT 791
Qy 781 TTACAATACAGCAATATAGCTGCTATCCAGATGATGATGATGATGATGATGATGAT 840
Db 792 TTACAATACAGCAATATAGCTGCTATCCAGATGATGATGATGATGATGATGATGAT 851
Qy 841 TGTGATGTTTCTAACGATTAAGTCAATGCGCAGTAGAGAGAGCCATCACTCAAGTAA 900
Db 852 TGTGATGTTTCTAACGATTAAGTCAATGCGCAGTAGAGAGAGCCATCACTCAAGTAA 911
Qy 901 TGTTCCTTCCGAAATTAACCAAGCAGTAGAACAACCCGCGACTGGGGA 949
Db 912 TGTTCCTTCCGAAATTAACCAAGCAGTAGAACAACCCGCGACTGGGGA 960

```

RESULT 13

PCT-US97-14436-139

Sequence 139, Application PC/TUS9714436

GENERAL INFORMATION:

```

APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nichols, Richard
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
TITLE OF INVENTION: POLYPEPTIDES AND THEIR USES
NUMBER OF SEQUENCES: 713
CORRESPONDENCE ADDRESS:
ADDRESS: Smithline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/14436
FILING DATE: 15-AUG-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,022
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50533
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:

```

INFORMATION FOR SEQ ID NO: 139:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 2172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

```

```

MOLECULE TYPE: Genomic DNA
PCT-US97-14436-139

```

```

Query Match 27.1%; Score 541; DB 1; Length 2172;
Best Local Similarity 99.5%; Pred. No. 1,1e-271;
Matches 811; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

```

Qy 1186 CCTAAATGGTCTAGCAATGAGTACCAAGTATATCTCACTCAATATGCCATTTCAGTAA 1245
Db 13 CCTAAATGGTCTAGCAATGAGTACCAAGTATATCTCACTCAATATGCCATTTCAGTAA 72
Qy 1246 CACAACCGAATCAGACAAAAAATATGAGCAAGTATGAAAGATGCTGCTTACGC 1305
Db 73 CACAACCGAATCAGACAAAAAATATGAGCAAGTATGAAAGATGCTGCTTACGC 132
Qy 1306 TGCCTTTGCAAAATGCTGGAATCTACTATTAACCAATGTATATCATTAAGTCTTTAG 1365
Db 133 TGCCTTTGCAAAATGCTGGAATCTACTATTAACCAATGTATATCATTAAGTCTTTAG 192
Qy 1366 TGAATGGAGTGAAGAAAGATGCTCTAATGTGGAACTGCTGCATGAGAAAGCAGAC 1425
Db 193 TGAATGGAGTGAAGAAAGATGCTCTAATGTGGAACTGCTGCATGAGAAAGCAGAC 252
Qy 1426 CTATATGATGACCGAGATGATGAAAAAGTCTTGAATGAGAACTGAGC- AAATGCTT 1484
Db 253 CTATATGATGACCGAGATGATGAAAAAGTCTTGAATGAGAACTGAGC- AAATGCTT 312
Qy 1485 ATCTTGCTTGCTCCCTCCTCAGGCTGCTTAACAGAGAACTCTAATCTATACAGAGAGAA 1544
Db 313 ATCTTGCTTGCTCCCTCCTCAGGCTGCTTAACAGAGAACTCTAATCTATACAGAGAGAA 372
Qy 1545 TTGAAAACCAATCAAGACCTCTCAATTTGTAGACCTGATGAACTATTTGCTGGCTATA 1604
Db 373 TTGAAAACCAATCAAGACCTCTCAATTTGTAGACCTGATGAACTATTTGCTGGCTATA 432
Qy 1605 CCGCTAAATATTCATGAGCTGTATGAGACAGGCTATTTCAACGCTGACACCACTTGTAG 1664
Db 433 CCGCTAAATATTCATGAGCTGTATGAGACAGGCTATTTCAACGCTGACACCACTTGTAG 492
Qy 1665 GCAATGGCTTACGCTGCTGCTGCAAAAGTTTACCGCTCTATGATGACCTTCTGTGAAG 1724
Db 493 GCAATGGCTTACGCTGCTGCTGCAAAAGTTTACCGCTCTATGATGACCTTCTGTGAAG 552
Qy 1725 GAAGCAATCCAGAAATGGAATATACAGAGGGGCTTACAGAAATGGAATTTGTAT 1784
Db 553 GAAGCAATCCAGAAATGGAATATACAGAGGGGCTTACAGAAATGGAATTTGTAT 612
Qy 1785 TTAATAATGCTGCTGCTTACGTGAGACTCACTGCTCAGACAAACCCCATCAACTG 1844
Db 613 TTAATAATGCTGCTGCTTACGTGAGACTCACTGCTCAGACAAACCCCATCAACTG 672
Qy 1845 AAAGTTCAAGCTCATCATGATGATGATTTCACTTCACTGCTGCTGCTGCTGCTGCTG 1904
Db 673 AAAGTTCAAGCTCATCATGATGATGATTTCACTTCACTGCTGCTGCTGCTGCTGCTG 732
Qy 1905 CAAATTAATGATGAGTACATCAATCCCTTAACATTAATGGAACAAATTAATTAACCCCG 1964
Db 733 CAAATTAATGATGAGTACATCAATCCCTTAACATTAATGGAACAAATTAATTAACCCCG 792
Qy 1965 ATCAACAAATCAGAAATCTCTCAACGACGACAAACA 1999
Db 793 ATCAACAAATCAGAAATCTCTCAACGACGACAAACA 827

```

RESULT 14

US-08-911-503-139

Sequence 139, Application US/08911503

```

GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael

```



```

? APPLICANT: Nicholas, Richard
? APPLICANT: Stodola, Robert
? TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
? TITLE OF INVENTION: POLYPEPTIDES AND THEIR USES
? NUMBER OF SEQUENCES: 713
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Smithline Beecham Corporation
? STREET: 709 Swedeland Road
? CITY: King of Prussia
? STATE: PA
? COUNTRY: USA
? ZIP: 19406-0939
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/911,503
? FILING DATE: 15-AUG-1997
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 60/024,022
? FILING DATE: 16-AUG-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Gimmil, Edward R
? REGISTRATION NUMBER: 38,891
? REFERENCE/DOCKET NUMBER: P50533
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 610-270-4478
? TELEFAX: 610-270-5090
? TELEX:
? INFORMATION FOR SEQ ID NO: 139:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2172 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: Genomic DNA
? US-08-911-503-139

Query Match      27.1%; Score 541; DB 13; Length 2172;
Best Local Similarity 99.5%; Pred. No. 1.1e-211;
Matches 811; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1186 CCTAATGCTAGAGATGCACTACCAAGATTCCTACTACTCAATTCAGTTCAAGTAA 1245
DB 13 CCTAAATGGTCTAGGAATGCACTACCAAGATTCCTACTACTCAATTCAGTTCAAGTAA 72
QY 1246 CACAACCGATTCAGACAAAAAATATGAGCAGATGAGTGAAGATGGCTGCTTACGC 1305
DB 73 CACAACCGATTCAGACAAAAAATATGAGCAGATGAGTGAAGATGGCTGCTTACGC 132
QY 1306 TGCCTTTGCAAAATGTTGAAGCTTACTATAAACAATGTATATCCATAAAGTCTTTAG 1365
DB 133 TGCCTTTGCAAAATGTTGAAGCTTACTATAAACAATGTATATCCATAAAGTCTTTAG 192
QY 1366 TGATGGAGTGAAGAAAGATTTCTATATGTCGGAACCTGTCCTGAAGAGAACGACAGC 1425
DB 193 TGATGGAGTGAAGAAAGATTTCTATATGTCGGAACCTGTCCTGAAGAGAACGACAGC 252
QY 1426 CTATATGATACGACATGATGTAAGAAAGTCTTACTATGAACTGAGC-NAATGCCT 1484
DB 253 CTATATGATACGACATGATGTAAGAAAGTCTTACTATGAACTGAGC-NAATGCCT 312
QY 1485 ATCTTGGTGTCCCTCAGAGCTGTGTAAGAACAGAACCTCTAATCTATACAGACAGAGAA 1544
DB 313 ATCTTGGTGTCCCTCAGAGCTGTGTAAGAACAGAACCTCTAATCTATACAGACAGAGAA 372
QY 1545 TTGAAAACCAATCAGACCTCTCAATTTGTAGCAGCTGATGAATCTTTTGTGGCTATA 1604
DB 373 TTGAAAACCAATCAGACCTCTCAATTTGTAGCAGCTGATGAATCTTTTGTGGCTATA 432

```

```

QY 1605 CGCCTAAATATTCATGCTGTATGACAGGCTATTTCAACCGTGTGACACCACTGTAG 1664
DB 433 CGCCTAAATATTCATGCTGTATGACAGGCTATTTCAACCGTGTGACACCACTGTAG 492
QY 1665 GCAATGCTTTAGCGTGCCTGCCAAAGTTTACCGCTCTATGATGACCTGCTGTGAAG 1724
DB 493 GCAATGCTTTAGCGTGCCTGCCAAAGTTTACCGCTCTATGATGACCTGCTGTGAAG 552
QY 1725 GAACATCCAGAGATTTGGAATATACACAGGGGCTCTACAGAAATGGAAATTCGTAT 1784
DB 553 GAACATCCAGAGATTTGGAATATACACAGGGGCTCTACAGAAATGGAAATTCGTAT 612
QY 1785 TTAATAATGGTCTGCTTCTACAGTGAAGTACCTGCTCCACACACCCCATCACTG 1844
DB 613 TTAATAATGGTCTGCTTCTACAGTGAAGTACCTGCTCCACACACCCCATCACTG 672
QY 1845 AAAGTTCAAGCTCATCATCAGATAGTTCAACTTACAGTCTAGCTCAACCACTCAAGCA 1904
DB 673 AAAGTTCAAGCTCATCATCAGATAGTTCAACTTACAGTCTAGCTCAACCACTCAAGCA 732
QY 1905 CAATATATGATGACATACCAATCTTAACATATATAGCAACATCAATCAACCCCTG 1964
DB 733 CAATATATGATGACATACCAATCTTAACATATATAGCAACATCAATCAACCCCTG 792
QY 1965 ATCACAAAAATCAGATCTCTCAACCAACCAACCA 1999
DB 793 ATCACAAAAATCAGATCTCTCAACCAACCAACCA 827

RESULT 15
US-08-911-503A-139
? Sequence 139, Application US/08911503A
? GENERAL INFORMATION:
? APPLICANT: Black, Michael
? APPLICANT: Hodgson, John
? APPLICANT: Knowles, David
? APPLICANT: Lonetto, Michael
? APPLICANT: Nicholas, Richard
? APPLICANT: Stodola, Robert
? TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
? TITLE OF INVENTION: POLYPEPTIDES AND THEIR USES
? NUMBER OF SEQUENCES: 713
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Smithline Beecham Corporation
? STREET: 709 Swedeland Road
? CITY: King of Prussia
? STATE: PA
? COUNTRY: USA
? ZIP: 19406-0939
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/911,503A
? FILING DATE: 15-AUG-1997
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 60/024,022
? FILING DATE: 16-AUG-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Gimmil, Edward R
? REGISTRATION NUMBER: 38,891
? REFERENCE/DOCKET NUMBER: P50533
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 610-270-4478
? TELEFAX: 610-270-5090
? TELEX:
? INFORMATION FOR SEQ ID NO: 139:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2172 base pairs

```



```

;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-911-503A-139

```

```

Query Match      27.1%; Score 541; DB 13; Length 2172;
Best Local Similarity 99.5%; Pred. No. 1.1e-271;
Matches 811; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

```

OY 1186 CCTAAATGCTAGAGATGAGTACCAAGTATTCCTACTACTCAAAATGCAATTCAGTAA 1245
    |||
DB 13 CCTAAATGCTAGAGATGAGTACCAAGTATTCCTACTACTCAAAATGCAATTCAGTAA 72
OY 1246 CCAACCCGATGAGCAAAAAATATGAGCAAGTACTGAAAAAGTGGCTGCTTACGC 1305
    |||
DB 73 CCAACCCGATGAGCAAAAAATATGAGCAAGTACTGAAAAAGTGGCTGCTTACGC 132
OY 1306 TGCCTTTGCAATGCTGAACTTACTATTAACCAATGATATTCATCAATGCTCTTTAG 1365
    |||
DB 133 TGCCTTTGCAATGCTGAACTTACTATTAACCAATGATATTCATCAATGCTCTTTAG 192
OY 1366 TGATGGAGTGAAGAAAGAGCTTCTTAATGTCGGAACCTGCGCATGAAGAAACGACAC 1425
    |||
DB 193 TGATGGAGTGAAGAAAGAGCTTCTTAATGTCGGAACCTGCGCATGAAGAAACGACAC 252
OY 1426 CTRATAGATGACCGACATATGAAAAACAGTCTTGACTTATGCACTGGACG-AAATGCT 1484
    |||
DB 253 CTRATAGATGACCGACATATGAAAAACAGTCTTGACTTATGCACTGGACG-AAATGCT 312
OY 1485 ATCTTGCTGGCTCCCTCGAGGCTGTAAGAAACGAGAACCTCTACTACTACAGAGAGAA 1544
    |||
DB 313 ATCTTGCTGGCTCCCTCGAGGCTGTAAGAAACGAGAACCTCTACTACTACAGAGAGAA 372
OY 1545 TTGAAACCCACATCAAGACCTCTCAATTTGTAGCACCTGATGAATTTGCTGGCTATA 1604
    |||
DB 373 TTGAAACCCACATCAAGACCTCTCAATTTGTAGCACCTGATGAATTTGCTGGCTATA 432
OY 1605 CCGCTAAATATTCATAGGCTGTATGAGACAGGCTATTTCAACGCTGACACCACTTTGAG 1664
    |||
DB 433 CCGCTAAATATTCATAGGCTGTATGAGACAGGCTATTTCAACGCTGACACCACTTTGAG 492
OY 1665 GCAATGCTGCTGAGGCTGCTGCAAAAGTTTACCGCTCTAATGATGACCTGCTGTAAG 1724
    |||
DB 493 GCAATGCTGCTGAGGCTGCTGCAAAAGTTTACCGCTCTAATGATGACCTGCTGTAAG 552
OY 1725 GAAGCAATCCAGAGATTTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAT 1784
    |||
DB 553 GAAGCAATCCAGAGATTTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAT 612
OY 1785 TTTAAAAATGCTGCTGTTTACGTGGAATCTCACCTGCTCCACAAACCCCATCAACTG 1844
    |||
DB 613 TTTAAAAATGCTGCTGTTTACGTGGAATCTCACCTGCTCCACAAACCCCATCAACTG 672
OY 1845 AAGAGTCAAGCTCATGATCAGATAGTGAATTCACAGTCTGCTCAACCATCCCAAGA 1904
    |||
DB 673 AAGAGTCAAGCTCATGATCAGATAGTGAATTCACAGTCTGCTCAACCATCCCAAGA 732
OY 1905 CAAATATATGATGAGTACTACCAATCTTAACATATATGCAAGAAATCAATTAACACCCCTG 1964
    |||
DB 733 CAAATATATGATGAGTACTACCAATCTTAACATATATGCAAGAAATCAATTAACACCCCTG 792
OY 1965 ATCAACAAAATAGAAATCTCTCAACGACGACAAACA 1999
    |||
DB 793 ATCAACAAAATAGAAATCTCTCAACGACGACAAACA 827

```

```

RESULT 16
US-60-068-175-596/c
; Sequence 596, Application US/60068175
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert E.

```

```

;
; APPLICANT: Corley, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Hann, Amy L.
; APPLICANT: Heath, Joe D.
; APPLICANT: Finney, Gregory L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE
; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 1175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/068,175
; FILING DATE: HERewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0009-2 P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-416
; INFORMATION FOR SEQ ID NO: 596:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9845 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: genomic DNA
; IMMEDIATE SOURCE:
; CLONE: SPN2c598
US-60-068-175-596

```

```

Query Match      14.7%; Score 293; DB 45; Length 9845;
Best Local Similarity 98.9%; Pred. No. 9.3e-142;
Matches 643; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

OY 1 TAAATCTACGACAAATAAATCACTCAATGCTGACTTGGGTTCTGAACGCGCGCTCAA 60
    |||
DB 9734 TAAATCTACGACAAATAAATCACTCAATGCTGACTTGGGTTCTGAACGCGCGCTCAA 9675
OY 61 TCGCCAGCTAATGATATTTCCACAGATTTGGTTAAGCAATGTTCTATGCAAGACCA 120
    |||
DB 9674 TCGCCAGCTAATGATATTTCCACAGATTTGGTTAAGCAATGTTCTATGCAAGACCA 9615
OY 121 TCGCTCTTCGACACAGGGGATGATACATCCGTATCCGTGGAGCTTTCTGGCAA 180
    |||
DB 9614 TCGCTCTTCGACACAGGGGATGATACATCCGTATCCGTGGAGCTTTCTGGCAA 9555
OY 181 TCTGCAAGCAATTCCTCCAGAGTGAATCACTCAACCCCAAGTATTAATTAATGAC 240
    |||
DB 9554 TCTGCAAGCAATTCCTCCAGAGTGAATCACTCAACCCCAAGTATTAATTAATGAC 9495
OY 241 TTACTTTTCACTGACTTCGACAGACTATTTCTCTGAAGGCTCAGGAAGCTTGTT 300
    |||
DB 9494 TTACTTTTCACTGACTTCGACAGACTATTTCTCTGAAGGCTCAGGAAGCTTGTT 9435
OY 301 AGCGATTGCTGATGAAACAAAGCAACGAAAGAAATCTTGACTACTATATAATAA 360
    |||
DB 9434 AGCGATTGCTGATGAAACAAAGCAACGAAAGAAATCTTGACTACTATATAATAA 9375
OY 361 GGTCTACATGCTAATGGAATGAGATGAGAGACAGAGCTCAAAATCTACTATGTTAA 420
    |||

```


Db 9374 GGTCTACATCTCTAANGGCACTATGGAATGCAGACAGCAGCTCAAAAACACTATGCTAA 9315
QY 421 AGACCTCAATATTTAAGTTAGTTACCTGAGCTTGGCTGGTGAATGCTCAGGCACC 480
Db 9314 AGACCTCAATATTTAAGTTAGTTACCTGAGCTTGGCTGGTGAATGCTCAGGCACC 9255
QY 481 AAACCAATATGACCCCTATTTCACATCCAGAGAGCCCAAGCCGCGAAACTTGCTCT 540
Db 9254 AAACCAATATGACCCCTATTTCACATCCAGAGAGCCCAAGCCGCGAAACTTGCTCT 9195
QY 541 ATCTGAATGAAAAAATCAAGGCTACATCTCTGCTGAACACTATGAGAAACAGCAATAC 600
Db 9194 ATCTGAATGAAAAAATCAAGGCTACATCTCTGCTGAACACTATGAGAAACAGCAATAC 9135
QY 601 ACCAATCTAGTGAAGTCAAGTCTCAATGCAAGTCAAGTCAATGCTG 650
Db 9134 ACCAATCTAGTGAAGTCAAGTCTCAATGCTCAAGTCAAGTCAATGCTG 9085

RESULT 17

US-60-029-960-934/C
; Sequence 934, Application US/60029960
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 1649
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/029,960
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 934:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 323 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-60-029-960-934

Query Match 11.2%; Score 223; DB 41; Length 323;
Best Local Similarity 100.0%; Pred. No. 4.5e-105;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1324 AACTTACTATAAACCAATGATATCCATAAGTCGCTTTAGTGGAGGTGAAAAAGA 1383
Db 293 AACTTACTATAAACCAATGATATCCATAAGTCGCTTTAGTGGAGGTGAAAAAGA 234
QY 1384 GTTCTCTAATGTGGAAGTGTGCGCATGAAGGAAGCAGCCATATGATGACCGCAT 1443
Db 233 GTTCTCTAATGTGGAAGTGTGCGCATGAAGGAAGCAGCCATATGATGACCGCAT 174

QY 1444 GATGAAAAACAGCTCTTACTATATGGAAGTGGAGCAAAATGCTATCTTGGCTCCCTCA 1503
Db 173 GATGAAAAACAGCTCTTACTATATGGAAGTGGAGCAAAATGCTATCTTGGCTCCCTCA 114
QY 1504 GGTGCTGTAACAGGAGCACTCTACTATACAGACGAGAAATT 1546
Db 113 GGTGCTGTAACAGGAGCACTCTACTATATACAGAGGAGAAATT 71

RESULT 18

US-09-583-110-1313/C
; Sequence 1313, Application US/09583110
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PAT000-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 1313
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-1313

Query Match 10.2%; Score 204; DB 22; Length 204;
Best Local Similarity 100.0%; Pred. No. 4e-95;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1762 CTACAGAAATGGAATGCTATTTAAAAATGCTGCTGCTTCTAGCTGGAAGTCACTGCG 1821
Db 204 CTACAGAAATGGAATGCTATTTAAAAATGCTGCTGCTTCTAGCTGGAAGTCACTGCG 145
QY 1822 TCCACAGCAACCCCATCACTGGAAGTTCAGCTCATCATGATGATGTTCAACTTCACA 1881
Db 144 TCCACAGCAACCCCATCACTGGAAGTTCAGCTCATCATGATGATGTTCAACTTCACA 85
QY 1882 GTCTAGCTCAACCACTCCAGCAAAATATATAGTACGATACCAATCTTAACATAATAC 1941
Db 84 GTCTAGCTCAACCACTCCAGCAAAATATATAGTACGATACCAATCTTAACATAATAC 25
QY 1942 GCAACAAATCAATACAAACCCCTGA 1965
Db 24 GCAACAAATCAATACAAACCCCTGA 1

RESULT 19

US-09-134-000-1534
; Sequence 1534, Application US/09134000A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS
; FILE REFERENCE: GTC-005
; CURRENT APPLICATION NUMBER: US/09/134,000A
; CURRENT FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 6810
; SEQ ID NO 1534
; LENGTH: 2370
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000-1534

Query Match 1.3%; Score 25; DB 15; Length 2370;

Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1226 TCAATGCCATTTCAGTAACAA 1250
Db 1495 tcaatgccatttcagtaacaa 1519

RESULT 20

US-09-070-927-217

; Sequence 217, Application US/09070927
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; APPLICANT: Patrick J. Dillon
; APPLICANT: Steven C. Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927
; FILING DATE: herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 217:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-070-927-217

Query Match 1.3%; Score 25; DB 14; Length 8395;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1226 TCAATGCCATTTCAGTAACAA 1250
Db 2589 TCAATGCCATTTCAGTAACAA 2613

RESULT 21
US-09-070-927A-217
; Sequence 217, Application US/09070927A
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; APPLICANT: Patrick J. Dillon
; APPLICANT: Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14

ATTORNEY/AGENT INFORMATION:

; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 217:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 8395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 217:

US-09-070-927A-217

Query Match 1.3%; Score 25; DB 14; Length 8395;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1226 TCAATGCCATTTCAGTAACAA 1250
Db 2589 TCAATGCCATTTCAGTAACAA 2613

RESULT 22

US-09-536-784-227

; Sequence 227, Application US/09536784
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997


```
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 227:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-536-784-227
SEQUENCE DESCRIPTION: SEQ ID NO: 227:

Query Match          1.2%; Score 23; DB 20; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TAAATCTAGACATATAAATC 23
Db 11 TAAATCTAGACATATAAATC 33

RESULT 23
US-09-765-271-227
; Sequence 227, Application US/09765271
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,271
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/536,784
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 227:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-765-271-227
SEQUENCE DESCRIPTION: SEQ ID NO: 227:

Query Match          1.2%; Score 23; DB 30; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TAAATCTAGACATATAAATC 23
Db 11 TAAATCTAGACATATAAATC 33

RESULT 24
US-09-765-272-227
; Sequence 227, Application US/09765272
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 227:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-765-272-227
SEQUENCE DESCRIPTION: SEQ ID NO: 227:

Query Match          1.2%; Score 23; DB 30; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TAAATCTAGACATATAAATC 23
Db 11 TAAATCTAGACATATAAATC 33

RESULT 25
US-09-045-649-108
; Sequence 108, Application US/60045649
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert E.
; APPLICANT: Corely, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Hann, Amy L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 1466
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
```


CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/045.649
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PM-0003P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 933 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
CLONE: EF1C0108
US-60-045-649-108

Query Match 1.1%; Score 22; DB 43; Length 933;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1229 AATGCCATTCAAGTAACACAA 1250
|||||
Db 708 AATGCCATTCAAGTAACACAA 729

RESULT 26
US-60-046-653-183
; Sequence 183, Application US/60046653
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert E.
; APPLICANT: Corley, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Hann, Amy L.
; APPLICANT: Heath, Joe D.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 1449
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/046.653
FILING DATE: HEREMITH
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PM-0003-1 P

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 183:
SEQUENCE CHARACTERISTICS:
LENGTH: 933 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
CLONE: EF1C183
US-60-046-653-183

Query Match 1.1%; Score 22; DB 43; Length 933;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1229 AATGCCATTCAAGTAACACAA 1250
|||||
Db 708 AATGCCATTCAAGTAACACAA 729

RESULT 27
PCT-US02-03987-6631
; Sequence 6631, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit
; FILE REFERENCE: ELITRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6631
; LENGTH: 2337
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2337)
PCT-US02-03987-6631

Query Match 1.1%; Score 22; DB 1; Length 2337;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1229 AATGCCATTCAAGTAACACAA 1250
|||||
Db 1465 aatgccattcaagtaacacaa 1486

RESULT 28
US-09-815-242-6631
; Sequence 6631, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815.242

APPLICANT: Kingsbury, Gillian A.
APPLICANT: Welch, Nadine S.
APPLICANT: McCarthy, Sean A.
APPLICANT: Williamson, Mark
APPLICANT: Richardson, Jennifer
APPLICANT: Macbeth, Kyle J.
APPLICANT: Fraser, Christopher C.
APPLICANT: Villevall, Jean-Luc M.G.
APPLICANT: Goodearl, Andrew D.J.
APPLICANT: Slios-Santiago, Inmaculada
APPLICANT: White, David
APPLICANT: Pan, Yang
APPLICANT: Busfield, Samantha J.
APPLICANT: Deeds, James
APPLICANT: Lee, John
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: 1600.1164-001
CURRENT APPLICATION NUMBER: US/09/649,163
CURRENT FILING DATE: 2000-08-25
PRIOR APPLICATION NUMBER: 60/150,608
PRIOR FILING DATE: 1999-08-25
NUMBER OF SEQ ID NOS: 10535
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2312
LENGTH: 537
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(537)
OTHER INFORMATION: n = A,T,C or G
US-09-649-163-2312

Query Match 1.1%; Score 21; DB 25; Length 537;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 572 GCTGAACAGTATGAGAAGCA 592
|||||
Db 250 GCTGAACAGTATGAGAAGCA 230

RESULT 32
US-09-652-916-1790/C
; Sequence 1790, Application US/09652916
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1164-001
; CURRENT APPLICATION NUMBER: US/09/652,916
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,103
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 10480
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1790
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(537)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-916-1790

Query Match 1.1%; Score 21; DB 25; Length 537;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 572 GCTGAACAGTATGAGAAGCA 592
|||||
Db 250 GCTGAACAGTATGAGAAGCA 230

RESULT 33
US-09-403-609A-25
; Sequence 25, Application US/09403609A
; GENERAL INFORMATION:
; APPLICANT: HAKENBECK, Regine
; TITLE OF INVENTION: DNA PROBES, METHOD AND KIT FOR IDENTIFYING
; TITLE OF INVENTION: ANTIBIOTIC-RESISTANT STRAINS OF BACTERIA
; FILE REFERENCE: 012627-011
; CURRENT APPLICATION NUMBER: US/09/403,609A
; CURRENT FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/DE98/01134
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: DE 197 17 346.2
; PRIOR FILING DATE: 1997-04-24
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-403-609A-25

Query Match 1.0%; Score 20; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1577 GCACCTGATGACTATTTCG 1596
|||||
Db 1 gcacctgacactatttgc 20

RESULT 34
US-09-535-896-44152
; Sequence 44152, Application US/09535896
; GENERAL INFORMATION:
; APPLICANT: Selhamer, Jeffrey J.
; APPLICANT: Deleageane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING HYDROLASES
; FILE REFERENCE: PD-1003 CIP
; CURRENT APPLICATION NUMBER: US/09/535,896
; CURRENT FILING DATE: 2000-03-24
; PRIOR APPLICATION data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 46268
; SOFTWARE: PERL Program
; SEQ ID NO 44152
; LENGTH: 242
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu01334414
US-09-535-896-44152

Query Match 1.0%; Score 20; DB 20; Length 242;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAAGCAACCAAGCAAGAA 338
|||||
Db 36 aaaagcaaccaagcaagaa 55


```
RESULT 35
US-09-515-128-12893/C
; Sequence 12893, Application US/09515128
; GENERAL INFORMATION:
; APPLICANT: Arteburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandi
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jessen, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Ralsi, Fariba
; APPLICANT: Smith, Benjamin
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Fel
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
; FILE REFERENCE: 728CIP
; CURRENT APPLICATION NUMBER: US/09/515,128
; CURRENT FILING DATE: 2000-02-28
; EARLIER APPLICATION NUMBER: 09/034,341
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12893
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(332)
; OTHER INFORMATION: n = A,T,C or G
US-09-515-128-12893

Query Match          1.0%; Score 20; DB 19; Length 332:
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1867 TACTCACTTCACAGCTCA 1886
      |||||||
Db 243 TACTCACTTCACAGCTCA 224
```

```
RESULT 36
US-08-688-870-855
; Sequence 855, Application US/08688870
; GENERAL INFORMATION:
; APPLICANT: Stuart, Susan G.
; APPLICANT: Delegeane, Angelo
; APPLICANT: Bills, Pamela K.
; APPLICANT: Pham, Mino Thu
; APPLICANT: Altus, Christina M.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Ingrid Erika Akerblom
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
```

```
; TITLE OF INVENTION: PERIPHERAL BLOOD MACROPHAGES
; NUMBER OF SEQUENCES: 2162
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1. for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/688,870
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Michael C. Cerrone
; REGISTRATION NUMBER: 39132
; REFERENCE/DOCKET NUMBER: PD-0040 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 855:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 201668
US-08-688-870-855
```

```
Query Match          1.0%; Score 20; DB 10; Length 355:
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1769 AATGAGAAATTCGATTTAA 1788
      |||||||
Db 186 AATGAGAAATTCGATTTAA 205
```

```
RESULT 37
US-60-001-754-855
; Sequence 855, Application US/60001754
; GENERAL INFORMATION:
; APPLICANT: Stuart, Susan G.
; APPLICANT: Delegeane, Angelo
; APPLICANT: Bills, Pamela K.
; APPLICANT: Pham, Mino Thu
; APPLICANT: Altus, Christina M.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Ingrid Erika Akerblom
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; NUMBER OF SEQUENCES: 2162
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3330 HILVIEW AVENUE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1. for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/001,754
```


FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: LUTHER, BARBARA J.
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PD-0040P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0195
TELEFAX: (415) 855-0555
INFORMATION FOR SEQ ID NO: 855:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 201668
US-60-001-734-855

Query Match 1.0%; Score 20; DB 39; Length 355;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1769 AATGAGAAATTCGTATTAA 1788
|||||
DB 186 AATGAGAAATTCGTATTAA 205

RESULT 38
US-09-205-070-44712
; Sequence 44712, Application US/09205070
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-748
; CURRENT APPLICATION NUMBER: US/09/205,070
; CURRENT FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 45207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44712
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-205-070-44712

Query Match 1.0%; Score 20; DB 16; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCACCAAGCAAGAAA 338
|||||
DB 275 aaaagcaccaagcaagaaa 294

RESULT 39
US-09-321-214-31801
; Sequence 31801, Application US/09321214
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Drmanac, Radoje
; APPLICANT: Ctkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Dmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Ford, John
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-740

; CURRENT APPLICATION NUMBER: US/09/321,214
; CURRENT FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: US 60/088,041
; EARLIER FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31801
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-321-214-31801

Query Match 1.0%; Score 20; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCACCAAGCAAGAAA 338
|||||
DB 275 aaaagcaccaagcaagaaa 294

RESULT 40
US-09-340-623-44712
; Sequence 44712, Application US/09340623
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-748CON1
; CURRENT APPLICATION NUMBER: US/09/340,623
; CURRENT FILING DATE: 1999-06-28
; EARLIER APPLICATION NUMBER: US 09/205,070
; EARLIER FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 45207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44712
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-340-623-44712

Query Match 1.0%; Score 20; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCACCAAGCAAGAAA 338
|||||
DB 275 aaaagcaccaagcaagaaa 294

RESULT 41
US-09-516-335-31801
; Sequence 31801, Application US/09516335
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brady
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jensen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroza, Mimi


```

? APPLICANT: Lomelli, Michelle
? APPLICANT: Nelken, Sarah
? APPLICANT: Nguyen, Kody
? APPLICANT: Nguyen, Lynne
? APPLICANT: Nguyen, Phuong
? APPLICANT: Nogra, Margie
? APPLICANT: Palencia, Servando
? APPLICANT: Raisi, Fariba
? APPLICANT: Randhwa, Gurpreet
? APPLICANT: Sidhu, Navjwan
? APPLICANT: Smith, Benjamin
? APPLICANT: Smythe, Ashleigh
? APPLICANT: Trian, Lien
? APPLICANT: Verna, Ron
? APPLICANT: Wachter, Adam
? APPLICANT: Wu, James
? TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
? TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
? FILE REFERENCE: 740CIP
? CURRENT APPLICATION NUMBER: US/09/516,335
? CURRENT FILING DATE: 2000-03-01
? EARLIER APPLICATION NUMBER: 09/321,214
? EARLIER FILING DATE: 1999-05-26
? EARLIER APPLICATION NUMBER: 60/088,041
? EARLIER FILING DATE: 1998-06-02
? NUMBER OF SEQ ID NOS: 31906
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 31801
? LENGTH: 371
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-516-335-31801

```

```

Query Match 1.0%; Score 20; DB 19; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 319 AAAAGCAACCAAGCAAGAA 338
Db 275 aaagcaaccaagcaagaaa 294

```

```

RESULT 42
US-09-733-811-31801
? Sequence 31801, Application US/09733811
? GENERAL INFORMATION:
? APPLICANT: Arterburn, Matthew
? APPLICANT: Asghari, Vida
? APPLICANT: Damavandi, Simin
? APPLICANT: Dickson, Mark
? APPLICANT: Drake, Jim
? APPLICANT: Drmanac, Radoje
? APPLICANT: Engleman, Carrie
? APPLICANT: Faulkner, Brandy
? APPLICANT: Fox, Melvin
? APPLICANT: Garcia, Veronica
? APPLICANT: Giedt, Gretchen
? APPLICANT: Jessen, Aaron
? APPLICANT: Jomek, Leni
? APPLICANT: Jones, Lee
? APPLICANT: Kita, David
? APPLICANT: Labat, Ivan
? APPLICANT: Laroya, Mimi
? APPLICANT: Lomelli, Michelle
? APPLICANT: Nelken, Sarah
? APPLICANT: Nguyen, Kody
? APPLICANT: Nguyen, Lynne
? APPLICANT: Nguyen, Phuong
? APPLICANT: Nogra, Margie
? APPLICANT: Palencia, Servando

```

```

? APPLICANT: Raisi, Fariba
? APPLICANT: Randhwa, Gurpreet
? APPLICANT: Sidhu, Navjwan
? APPLICANT: Smith, Benjamin
? APPLICANT: Smythe, Ashleigh
? APPLICANT: Tkach, Joe
? TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
? TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
? FILE REFERENCE: 740CIP
? CURRENT APPLICATION NUMBER: US/09/733,811
? CURRENT FILING DATE: 2000-12-08
? PRIOR APPLICATION NUMBER: 09/321,214
? PRIOR FILING DATE: 1999-05-26
? PRIOR APPLICATION NUMBER: 60/088,041
? PRIOR FILING DATE: 1998-06-02
? NUMBER OF SEQ ID NOS: 31906
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 31801
? LENGTH: 371
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-733-811-31801

```

```

Query Match 1.0%; Score 20; DB 29; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 319 AAAAGCAACCAAGCAAGAA 338
Db 275 aaagcaaccaagcaagaaa 294

```

```

RESULT 43
US-09-733-811A-31801
? Sequence 31801, Application US/09733811A
? GENERAL INFORMATION:
? APPLICANT: Hyseq, Inc.
? APPLICANT: Drmanac, Radoje
? APPLICANT: Crkvenjakov, Radomir
? APPLICANT: Dickson, Mark
? APPLICANT: Drmanac, Snezana
? APPLICANT: Labat, Ivan
? APPLICANT: Leshkowitz, Dena
? APPLICANT: Kita, David
? APPLICANT: Ford, John
? TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
? TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
? FILE REFERENCE: 20411-740
? CURRENT APPLICATION NUMBER: US/09/733,811A
? CURRENT FILING DATE: 2000-12-08
? PRIOR APPLICATION NUMBER: 09/321,214
? PRIOR FILING DATE: 1999-05-26
? NUMBER OF SEQ ID NOS: 31906
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 31801
? LENGTH: 371
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-733-811A-31801

```

```

Query Match 1.0%; Score 20; DB 29; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 319 AAAAGCAACCAAGCAAGAA 338
Db 275 aaagcaaccaagcaagaaa 294

```

```

RESULT 44
US-09-898-888-44712

```



```
; Sequence 44712, Application US/09898888
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-748CON1
; CURRENT APPLICATION NUMBER: US/09/898,888
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 09/340,623
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 45207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44712
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-898-888-44712
```

```
Query Match 1.0%; Score 20; DB 33; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 319 AAAAGCACCAAGCAGAGAA 338
|||||
```

```
Db 275 aaagcaaccaagcaagaa 294
```

```
RESULT 45
US-09-898-888A-44712
; Sequence 44712, Application US/09898888A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-748CON1
; CURRENT APPLICATION NUMBER: US/09/898,888A
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/340,623
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/205,070
; PRIOR FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 45207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44712
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-898-888A-44712
```

```
Query Match 1.0%; Score 20; DB 33; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 319 AAAAGCACCAAGCAGAGAA 338
|||||
```

```
Db 275 aaagcaaccaagcaagaa 294
```

```
RESULT 46
US-09-975-640-31801
; Sequence 31801, Application US/09975640
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asgnari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
```

```
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Melken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjivan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/975,640
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/733,811
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/088,041
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31801
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-975-640-31801
```

```
Query Match 1.0%; Score 20; DB 36; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 319 AAAAGCACCAAGCAGAGAA 338
|||||
```

```
Db 275 aaagcaaccaagcaagaa 294
```

```
RESULT 47
US-09-975-640A-31801
; Sequence 31801, Application US/09975640A
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asgnari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
```



```

; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kodj
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogira, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Ralski, Farba
; APPLICANT: Randwa, Gurpreet
; APPLICANT: Sidhu, Navjivan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verma, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/975,640A
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/733,811
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/088,041
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31801
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-975-640A-31801
```

```

Query Match 1.0%; Score 20; DB 36; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 319 AAAAGCACCACGACGAAA 338
Db 275 aaaagcaaccagcaagaata 294
```

```

RESULT 48
US-09-271-490-13695/c
; Sequence 13695, Application US/09271490
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-767
; CURRENT APPLICATION NUMBER: US/09/271,490
; CURRENT FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 19424
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13695
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(378)
; OTHER INFORMATION: n = A,T,C or G
US-09-271-490-13695
```

```

Query Match 1.0%; Score 20; DB 16; Length 378;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 1867 TAGTTCAACTTCACAGCTCTA 1886
Db 131 TAGTTCAACTTCACAGCTCTA 112
```

```

RESULT 49
US-09-925-552-13695/c
; Sequence 13695, Application US/09925552
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-767
; CURRENT APPLICATION NUMBER: US/09/925,552
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/271,490
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 19424
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13695
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(378)
; OTHER INFORMATION: n = A,T,C or G
US-09-925-552-13695
```

```

Query Match 1.0%; Score 20; DB 34; Length 378;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 1867 TAGTTCAACTTCACAGCTCTA 1886
Db 131 TAGTTCAACTTCACAGCTCTA 112
```

```

RESULT 50
US-10-032-354-13695/c
; Sequence 13695, Application US/10032354
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-767
; CURRENT APPLICATION NUMBER: US/10/032,354
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/271,490
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 19424
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13695
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(378)
; OTHER INFORMATION: n = A,T,C or G
US-10-032-354-13695
```

```

Query Match 1.0%; Score 20; DB 37; Length 378;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 1867 TAGTTCAACTTCACAGCTCTA 1886
Db 131 TAGTTCAACTTCACAGCTCTA 112
```


Search completed: June 13, 2002, 13:40:40
Job time: 9593 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 11:34:23 ; Search time 277.73 Seconds
(without alignments)
10116.359 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATCTACGACATATAA.....ATCCTCAACGACGACACCA 1999

Scoring table: OLIGO_NMC
Gapop 60.0 , Gapext 60.0

Searched: 997330 seqs, 70275440 residues

Word size : 15

Total number of hits satisfying chosen parameters: 2617

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCR_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1999	100.0	1999	4	US-08-961-083-1
2	23	1.2	33	4	US-08-961-083-227
3	21	1.1	1389	5	US-09-540-209B-4827
4	20	1.0	332	5	US-09-721-544-12893
5	20	1.0	502	6	US-10-027-632-240647
6	20	1.0	502	6	US-10-027-632-240648
7	20	1.0	585	6	US-10-027-632-214358
8	20	1.0	4508	5	US-09-053-375B-301
9	19	1.0	427	5	US-09-620-393B-634
10	19	1.0	428	5	US-09-620-393B-8776
11	19	1.0	507	6	US-10-027-632-98352
12	19	1.0	1079	6	US-10-027-632-250559
13	19	1.0	1079	6	US-10-027-632-250570
14	19	1.0	1299	7	US-09-360-039-23856
15	19	1.0	1299	7	US-09-360-039-24027
16	19	1.0	1299	7	US-09-360-039-42540
17	19	1.0	1920	7	US-09-539-331D-25958
18	18	0.9	166	5	US-09-539-800C-2484
19	18	0.9	229	5	US-09-540-210B-30401
20	18	0.9	250	5	US-09-539-331D-33884
21	18	0.9	255	5	US-09-539-800C-15921
22	18	0.9	264	5	US-09-540-210B-24058
23	18	0.9	268	5	US-09-540-210B-24058
24	18	0.9	277	5	US-09-919-002-7641
25	18	0.9	285	5	US-09-539-331D-21402
26	18	0.9	285	5	US-09-539-331D-21402

27	18	0.9	296	5	US-09-539-331D-34826	Sequence 34826, A
28	18	0.9	457	6	US-10-027-632-94603	Sequence 94603, A
29	18	0.9	457	6	US-10-027-632-94604	Sequence 94604, A
30	18	0.9	457	6	US-10-027-632-305498	Sequence 305498, A
31	18	0.9	457	6	US-10-027-632-305499	Sequence 305499, A
32	18	0.9	457	6	US-10-027-632-305500	Sequence 305500, A
33	18	0.9	478	6	US-10-027-632-286371	Sequence 286371, A
34	18	0.9	512	6	US-10-097-105-1069	Sequence 1069, Ap
35	18	0.9	526	5	US-09-620-393B-6548	Sequence 6548, Ap
36	18	0.9	532	6	US-10-027-632-153305	Sequence 153305, A
37	18	0.9	581	6	US-10-027-632-81926	Sequence 81926, A
38	18	0.9	850	6	US-10-027-632-157285	Sequence 157285, A
39	18	0.9	1229	6	US-10-106-698-791	Sequence 791, App
40	18	0.9	1281	5	US-09-935-625-18639	Sequence 18639, A
41	18	0.9	1293	7	US-09-935-625-27058	Sequence 27058, A
42	18	0.9	1293	7	US-09-935-625-47071	Sequence 47071, A
43	18	0.9	1293	7	US-09-935-625-47071	Sequence 47071, A
44	18	0.9	1328	5	US-09-919-002-13066	Sequence 13066, A
45	18	0.9	1555	6	US-10-027-632-253001	Sequence 253001, A
46	18	0.9	1555	6	US-10-027-632-253002	Sequence 253002, A
47	18	0.9	1555	6	US-10-027-632-253003	Sequence 253003, A
48	18	0.9	1555	6	US-10-027-632-253004	Sequence 253004, A
49	18	0.9	1773	7	US-09-360-039-24796	Sequence 24796, A
50	18	0.9	2288	6	US-10-115-123-53	Sequence 53, App1
51	18	0.9	2773	6	US-10-027-632-111835	Sequence 111835, A
52	18	0.9	3201	1	PCR-US02-13142-2190	Sequence 2190, Ap
53	18	0.9	3201	1	PCR-US02-13142-2190	Sequence 2190, Ap
54	18	0.9	3201	6	US-10-128-714-2190	Sequence 2190, Ap
55	18	0.9	3201	6	US-10-128-714-2190	Sequence 2190, Ap
56	18	0.9	3293	1	PCR-US02-13142-1190	Sequence 1190, Ap
57	18	0.9	3293	1	PCR-US02-13142-1190	Sequence 1190, Ap
58	18	0.9	3293	6	US-10-128-714-1190	Sequence 1190, Ap
59	18	0.9	3293	6	US-10-128-714-1190	Sequence 1190, Ap
60	18	0.9	5293	1	PCR-US02-13142-5190	Sequence 5190, Ap
61	18	0.9	5293	1	PCR-US02-13142-5190	Sequence 5190, Ap
62	18	0.9	5293	6	US-10-128-714-15190	Sequence 15190, Ap
63	18	0.9	5293	6	US-10-128-714-15190	Sequence 15190, Ap
64	18	0.9	13165	6	US-10-105-299-9019	Sequence 9019, Ap
65	18	0.9	13165	6	US-10-105-299-9019	Sequence 9019, Ap
66	18	0.9	22008	1	PCR-US02-09239-242	Sequence 242, App
67	18	0.9	22008	1	PCR-US02-09188-900	Sequence 900, App
68	18	0.9	22008	1	PCR-US02-09257-581	Sequence 581, App
69	18	0.9	22008	1	PCR-US02-09370-930	Sequence 930, App
70	18	0.9	22008	1	PCR-US02-09922-568	Sequence 568, App
71	18	0.9	22008	6	US-10-105-299-6302	Sequence 6302, Ap
72	18	0.9	713059	6	US-10-027-632-174581	Sequence 174581, A
73	17	0.9	27	4	US-08-961-083-228	Sequence 228, App
74	17	0.9	220	6	US-10-027-632-70744	Sequence 70744, A
75	17	0.9	250	5	US-09-539-331D-7293	Sequence 7293, Ap
76	17	0.9	252	5	US-09-975-254-17107	Sequence 17107, A
77	17	0.9	253	5	US-09-975-254-17107	Sequence 17107, A
78	17	0.9	253	5	US-09-460-592B-136	Sequence 136, App
79	17	0.9	264	6	US-10-027-632-72513	Sequence 72513, A
80	17	0.9	276	5	US-09-540-210B-4006	Sequence 34006, A
81	17	0.9	382	5	US-09-721-544-13314	Sequence 13314, A
82	17	0.9	385	6	US-10-027-632-39637	Sequence 39637, A
83	17	0.9	402	6	US-10-099-926-3330	Sequence 330, App
84	17	0.9	437	5	US-09-539-331D-29344	Sequence 29344, A
85	17	0.9	443	6	US-10-027-632-189682	Sequence 189682, A
86	17	0.9	457	6	US-10-027-632-79353	Sequence 79353, A
87	17	0.9	457	6	US-10-027-632-79354	Sequence 79354, A
88	17	0.9	457	6	US-10-027-632-301173	Sequence 301173, A
89	17	0.9	457	6	US-10-027-632-301173	Sequence 301173, A
90	17	0.9	472	6	US-10-027-632-89245	Sequence 89245, A
91	17	0.9	546	6	US-10-027-632-258977	Sequence 258977, A
92	17	0.9	546	6	US-10-027-632-28978	Sequence 28978, A
93	17	0.9	558	6	US-10-027-632-89473	Sequence 89473, A
94	17	0.9	558	6	US-10-027-632-89473	Sequence 89473, A
95	17	0.9	558	6	US-10-027-632-89473	Sequence 89473, A
96	17	0.9	565	6	US-10-027-632-262082	Sequence 262082, A
97	17	0.9	573	6	US-10-027-632-26318	Sequence 26318, A
98	17	0.9	623	6	US-10-027-632-300	Sequence 300, App
99	17	0.9	623	6	US-10-027-632-345	Sequence 345, App

c 100	17	0.9	633	6	US-10-027-632-54574	Sequence 54574, A	c 173	16	0.8	386	6	US-10-125-968-905	Sequence 905, App
c 101	17	0.9	633	6	US-10-027-632-293998	Sequence 293998, A	c 174	16	0.8	386	6	US-10-027-632-178887	Sequence 178887, A
c 102	17	0.9	642	6	US-10-027-632-256087	Sequence 256087, A	c 175	16	0.8	395	6	US-10-116-712-634	Sequence 634, App
c 103	17	0.9	643	6	US-10-027-632-211646	Sequence 211646, A	c 176	16	0.8	397	6	US-10-099-926-1717	Sequence 1717, App
c 104	17	0.9	643	6	US-10-027-632-211647	Sequence 211647, A	c 177	16	0.8	404	5	US-09-721-544-1213	Sequence 1213, App
c 105	17	0.9	643	6	US-10-027-632-211648	Sequence 211648, A	c 178	16	0.8	425	5	US-09-721-544-22282	Sequence 22282, App
c 106	17	0.9	644	6	US-10-027-632-192473	Sequence 192473, A	c 179	16	0.8	430	6	US-10-027-632-294776	Sequence 294776, A
c 107	17	0.9	662	6	US-10-027-632-127685	Sequence 127685, A	c 180	16	0.8	431	5	US-09-919-002-7411	Sequence 7411, App
c 108	17	0.9	711	6	US-10-027-632-22033	Sequence 22033, A	c 181	16	0.8	432	6	US-10-027-632-308516	Sequence 308516, A
c 109	17	0.9	720	6	US-10-027-632-23633	Sequence 23633, A	c 182	16	0.8	432	6	US-10-027-632-181299	Sequence 181299, A
c 110	17	0.9	734	6	US-10-027-632-11064	Sequence 11064, A	c 183	16	0.8	437	6	US-10-027-632-94829	Sequence 94829, A
c 111	17	0.9	734	6	US-10-027-632-11064	Sequence 11064, A	c 184	16	0.8	460	6	US-10-027-632-65703	Sequence 65703, A
c 112	17	0.9	776	6	US-10-027-632-168427	Sequence 168427, A	c 185	16	0.8	462	5	US-09-685-404A-203	Sequence 203, App
c 113	17	0.9	779	6	US-10-027-632-145394	Sequence 145394, A	c 186	16	0.8	464	5	US-09-620-393B-6694	Sequence 207, App
c 114	17	0.9	975	7	US-09-360-039-40220	Sequence 40220, A	c 187	16	0.8	467	5	US-09-685-404A-207	Sequence 207, App
c 115	17	0.9	1058	5	US-09-919-002-5787	Sequence 5787, App	c 188	16	0.8	468	6	US-10-027-632-189086	Sequence 189086, A
c 116	17	0.9	1202	5	US-09-882-227-509	Sequence 509, App	c 189	16	0.8	468	6	US-10-027-632-189087	Sequence 189087, A
c 117	17	0.9	1244	5	US-09-919-002-1145	Sequence 1145, App	c 190	16	0.8	470	6	US-10-027-632-84940	Sequence 84940, A
c 118	17	0.9	1244	5	US-09-919-002-2529	Sequence 2529, App	c 191	16	0.8	470	6	US-10-027-632-178926	Sequence 178926, A
c 119	17	0.9	1284	7	US-09-130-080-3	Sequence 47007, A	c 192	16	0.8	471	6	US-10-027-632-82965	Sequence 82965, A
c 120	17	0.9	1284	7	US-09-130-080-3	Sequence 47007, A	c 193	16	0.8	471	6	US-10-027-632-82966	Sequence 82966, A
c 121	17	0.9	1531	6	US-10-115-123-32	Sequence 32, App1	c 194	16	0.8	473	5	US-09-620-393B-5469	Sequence 5469, App
c 122	17	0.9	1531	6	US-10-115-123-32	Sequence 32, App1	c 195	16	0.8	476	6	US-10-027-632-50656	Sequence 50656, App
c 123	17	0.9	1906	5	US-09-442-384B-599	Sequence 599, App	c 196	16	0.8	480	6	US-10-027-632-41506	Sequence 41506, App
c 124	17	0.9	2013	6	US-10-106-698-1619	Sequence 1619, App	c 197	16	0.8	484	5	US-09-721-544-5673	Sequence 5673, App
c 125	17	0.9	2166	7	US-09-360-039-36422	Sequence 36422, A	c 198	16	0.8	484	7	US-09-377-240-7841	Sequence 7841, App
c 126	17	0.9	2346	6	US-10-027-632-110029	Sequence 110029, A	c 199	16	0.8	486	5	US-09-721-544-5672	Sequence 5672, App
c 127	17	0.9	2497	7	US-09-360-039-26849	Sequence 26849, A	c 200	16	0.8	486	5	US-09-721-544-2085	Sequence 2085, App
c 128	17	0.9	3247	6	US-10-146-473-34	Sequence 34, App1	c 201	16	0.8	484	6	US-10-027-632-69908	Sequence 69908, A
c 129	17	0.9	3510	7	US-09-360-039-24782	Sequence 24782, A	c 202	16	0.8	507	6	US-10-027-632-34500	Sequence 34500, A
c 130	17	0.9	22680	6	US-10-105-299-11585	Sequence 11585, A	c 203	16	0.8	513	6	US-10-027-632-188432	Sequence 188432, A
c 131	17	0.9	54863	6	US-10-105-299-8039	Sequence 8039, App	c 204	16	0.8	519	6	US-10-027-632-6820	Sequence 6820, App
c 132	17	0.9	54877	1	PCT-US02-09188-1314	Sequence 1314, App	c 205	16	0.8	524	6	US-10-027-632-312273	Sequence 312273, A
c 133	17	0.9	54877	1	PCT-US02-09370-1372	Sequence 1372, App	c 206	16	0.8	527	6	US-10-027-632-6809	Sequence 6809, App
c 134	17	0.9	54877	1	PCT-US02-09992-863	Sequence 863, App	c 207	16	0.8	527	6	US-10-027-632-323318	Sequence 323318, A
c 135	17	0.9	54877	6	US-10-105-299-6896	Sequence 6896, App	c 208	16	0.8	529	5	US-09-721-544-1228	Sequence 1228, App
c 136	17	0.9	54877	6	US-10-105-299-8040	Sequence 8040, App	c 209	16	0.8	530	5	US-09-620-393B-8742	Sequence 8742, App
c 137	17	0.9	120344	6	US-10-105-299-9867	Sequence 9867, App	c 210	16	0.8	532	6	US-10-027-632-58115	Sequence 58115, A
c 138	17	0.9	160820	6	PCT-US02-11086-5	Sequence 5, App11	c 211	16	0.8	542	6	US-10-027-632-59136	Sequence 59136, A
c 139	17	0.9	2940917	6	US-10-027-632-174763	Sequence 174763, A	c 212	16	0.8	549	5	US-09-991-936-1290	Sequence 936, App
c 140	16	0.8	20	6	US-10-130-080-55	Sequence 55, App1	c 213	16	0.8	550	7	US-09-377-240-5728	Sequence 5728, App
c 141	16	0.8	168	5	US-09-539-331D-17706	Sequence 17706, A	c 214	16	0.8	554	6	US-10-027-632-281424	Sequence 281424, A
c 142	16	0.8	180	5	US-09-539-331D-4063	Sequence 4063, App	c 215	16	0.8	557	6	US-10-027-632-77645	Sequence 77645, A
c 143	16	0.8	188	6	US-10-027-632-53214	Sequence 53214, A	c 216	16	0.8	557	6	US-10-027-632-300549	Sequence 300549, A
c 144	16	0.8	208	5	US-09-539-800C-5010	Sequence 5010, App	c 217	16	0.8	551	6	US-10-027-632-134117	Sequence 134117, A
c 145	16	0.8	216	5	US-09-899-575-80	Sequence 80, App1	c 218	16	0.8	571	6	US-10-027-632-215927	Sequence 215927, A
c 146	16	0.8	239	5	US-09-540-210B-29605	Sequence 29605, A	c 219	16	0.8	572	6	US-10-027-632-71487	Sequence 71487, A
c 147	16	0.8	245	5	US-09-539-331D-21787	Sequence 21787, A	c 220	16	0.8	572	6	US-10-027-632-293800	Sequence 293800, A
c 148	16	0.8	246	5	US-09-539-331D-27254	Sequence 27254, A	c 221	16	0.8	573	6	US-10-027-632-235623	Sequence 235623, A
c 149	16	0.8	247	5	US-09-539-806B-1690	Sequence 1690, App	c 222	16	0.8	577	6	US-10-105-299-9026	Sequence 9026, App
c 150	16	0.8	248	5	US-09-539-800C-9055	Sequence 9055, App	c 223	16	0.8	586	6	US-10-027-632-258363	Sequence 258363, A
c 151	16	0.8	253	5	US-09-975-254-26610	Sequence 26610, A	c 224	16	0.8	587	6	US-10-027-632-107979	Sequence 107979, A
c 152	16	0.8	259	5	US-09-539-800C-2513	Sequence 2513, App	c 225	16	0.8	587	6	US-10-027-632-170780	Sequence 170780, A
c 153	16	0.8	264	5	US-09-975-254-29322	Sequence 29322, A	c 226	16	0.8	587	6	US-10-027-632-47971	Sequence 47971, A
c 154	16	0.8	266	6	US-10-027-632-270234	Sequence 270234, A	c 227	16	0.8	587	6	US-10-027-632-47972	Sequence 47972, A
c 155	16	0.8	270	6	US-10-099-926-1640	Sequence 1640, App	c 228	16	0.8	588	6	US-10-027-632-44003	Sequence 44003, A
c 156	16	0.8	274	5	US-09-539-331D-24963	Sequence 24963, A	c 229	16	0.8	588	6	US-10-027-632-44004	Sequence 44004, A
c 157	16	0.8	276	5	US-09-540-210B-26442	Sequence 26442, A	c 230	16	0.8	588	6	US-10-027-632-44005	Sequence 44005, A
c 158	16	0.8	287	6	US-10-099-926-1718	Sequence 1718, App	c 231	16	0.8	588	6	US-10-027-632-44006	Sequence 44006, A
c 159	16	0.8	292	6	US-10-027-632-65077	Sequence 65077, A	c 232	16	0.8	588	6	US-10-027-632-30096	Sequence 30096, A
c 160	16	0.8	309	5	US-09-899-575-93	Sequence 93, App1	c 233	16	0.8	589	6	US-10-027-632-42858	Sequence 42858, A
c 161	16	0.8	311	5	US-09-539-800C-9580	Sequence 9580, App	c 234	16	0.8	589	6	US-10-027-632-42859	Sequence 42859, A
c 162	16	0.8	314	5	US-09-721-544-16447	Sequence 16447, A	c 235	16	0.8	609	6	US-10-027-632-107012	Sequence 107012, A
c 163	16	0.8	315	5	US-09-721-544-20031	Sequence 20031, A	c 236	16	0.8	609	6	US-10-027-632-254312	Sequence 254312, A
c 164	16	0.8	335	6	US-10-027-632-53601	Sequence 53601, A	c 237	16	0.8	612	6	US-10-027-632-280430	Sequence 280430, A
c 165	16	0.8	337	5	US-09-721-544-14946	Sequence 14946, A	c 238	16	0.8	612	6	US-10-027-632-280431	Sequence 280431, A
c 166	16	0.8	353	5	US-09-721-544-10664	Sequence 10664, A	c 239	16	0.8	618	6	US-10-027-632-101574	Sequence 101574, A
c 167	16	0.8	354	5	US-09-721-544-13294	Sequence 13294, A	c 240	16	0.8	618	6	US-10-027-632-101575	Sequence 101575, A
c 168	16	0.8	364	5	US-09-721-544-17222	Sequence 17222, A	c 241	16	0.8	623	7	US-10-027-632-197807	Sequence 197807, A
c 169	16	0.8	367	1	PCT-US02-12611-186	Sequence 186, App	c 242	16	0.8	623	7	US-09-377-240-3511	Sequence 3511, App
c 170	16	0.8	367	6	US-10-125-968-186	Sequence 186, App	c 243	16	0.8	625	6	US-10-027-632-84363	Sequence 84363, A
c 171	16	0.8	373	5	US-09-721-544-18631	Sequence 18631, A	c 244	16	0.8	625	6	US-10-027-632-269318	Sequence 269318, A
c 172	16	0.8	386	1	PCT-US02-12612-905	Sequence 905, App	c 245	16	0.8	625	6	US-10-027-632-269319	Sequence 269319, A

C 246	16	0.8	626	6	US-10-027-632-78176	
C 247	16	0.8	630	6	US-10-027-632-27921	
C 248	16	0.8	630	6	US-10-027-632-308413	
C 249	16	0.8	637	6	US-10-027-632-218502	
C 250	16	0.8	637	6	US-10-027-632-228230	
C 251	16	0.8	639	6	US-10-027-632-285397	
C 252	16	0.8	643	6	US-10-027-632-267586	
C 253	16	0.8	650	6	US-10-027-632-188190	
C 254	16	0.8	650	6	US-10-027-632-231937	
C 255	16	0.8	650	6	US-10-027-632-126527	
C 256	16	0.8	651	6	US-10-027-632-126528	
C 257	16	0.8	654	6	US-10-027-632-44023	
C 258	16	0.8	657	6	US-10-027-632-101477	
C 259	16	0.8	659	6	US-10-027-632-22811	
C 260	16	0.8	659	6	US-10-027-632-22812	
C 261	16	0.8	660	6	US-10-027-632-2075	
C 262	16	0.8	660	6	US-10-027-632-2075	
C 263	16	0.8	661	6	US-10-027-632-248483	
C 264	16	0.8	664	6	US-10-027-632-204053	
C 265	16	0.8	664	6	US-10-027-632-204054	
C 266	16	0.8	668	6	US-10-027-632-7191	
C 267	16	0.8	668	6	US-10-027-632-224984	
C 268	16	0.8	668	6	US-10-027-632-228470	
C 269	16	0.8	668	6	US-10-027-632-228471	
C 270	16	0.8	668	6	US-10-027-632-228472	
C 271	16	0.8	668	6	US-10-027-632-228473	
C 272	16	0.8	668	6	US-10-027-632-228474	
C 273	16	0.8	668	6	US-10-027-632-228475	
C 274	16	0.8	668	6	US-10-027-632-151638	
C 275	16	0.8	668	6	US-10-027-632-151639	
C 276	16	0.8	668	6	US-10-027-632-151640	
C 277	16	0.8	668	6	US-10-027-632-151641	
C 278	16	0.8	697	6	US-10-027-632-32430	
C 279	16	0.8	713	6	US-10-027-632-25909	
C 280	16	0.8	718	6	US-10-027-632-164051	
C 281	16	0.8	718	6	US-10-027-632-164052	
C 282	16	0.8	722	6	US-10-027-632-151457	
C 283	16	0.8	725	6	US-09-634-754C-803	
C 284	16	0.8	732	6	US-10-027-632-33179	
C 285	16	0.8	737	6	US-10-027-632-28056	
C 286	16	0.8	737	6	US-10-027-632-28057	
C 287	16	0.8	737	6	US-10-027-632-98623	
C 288	16	0.8	739	6	US-10-027-632-110652	
C 289	16	0.8	760	6	US-10-027-632-119220	
C 290	16	0.8	773	6	US-10-027-632-29955	
C 291	16	0.8	774	6	US-10-027-632-12339	
C 292	16	0.8	780	6	US-10-027-632-162813	
C 293	16	0.8	785	6	US-10-027-632-174200	
C 294	16	0.8	785	6	US-10-027-632-174201	
C 295	16	0.8	789	6	US-10-027-632-156257	
C 296	16	0.8	801	6	US-10-027-632-7134	
C 297	16	0.8	812	6	US-10-027-632-156402	
C 298	16	0.8	812	6	US-10-027-632-156403	
C 299	16	0.8	816	6	US-10-027-632-156404	
C 300	16	0.8	816	6	US-10-027-632-170387	
C 301	16	0.8	816	6	US-10-027-632-170388	
C 302	16	0.8	816	6	US-10-027-632-170389	
C 303	16	0.8	824	6	US-10-027-632-157911	
C 304	16	0.8	824	6	US-10-027-632-157912	
C 305	16	0.8	826	6	US-10-027-632-134116	
C 306	16	0.8	826	6	US-10-027-632-134116	
C 307	16	0.8	856	6	US-10-027-632-120426	
C 308	16	0.8	856	6	US-10-027-632-120427	
C 309	16	0.8	863	6	US-10-027-632-167880	
C 310	16	0.8	866	6	US-10-027-632-167880	
C 311	16	0.8	873	7	US-60-360-039-46819	
C 312	16	0.8	874	7	US-60-360-039-37207	
C 313	16	0.8	884	5	US-09-919-002-4009	
C 314	16	0.8	907	6	US-10-027-632-153864	
C 315	16	0.8	907	6	US-10-027-632-153865	
C 316	16	0.8	907	6	US-10-027-632-153866	
C 317	16	0.8	947	5	US-09-634-754C-110	
C 318	16	0.8	970	6	US-10-027-632-100933	

Sequence 78176, A	319	16	0.8	991	6	US-10-027-632-252940
Sequence 27921, A	320	16	0.8	994	6	US-10-027-632-31531
Sequence 308413, A	C 321	16	0.8	1008	6	US-10-027-632-120200
Sequence 218502, A	C 322	16	0.8	1008	6	US-10-027-632-120201
Sequence 22930, A	C 323	16	0.8	1052	6	US-10-027-632-9749
Sequence 228597, A	C 324	16	0.8	1052	6	US-10-027-632-9750
Sequence 267586, A	C 325	16	0.8	1055	5	US-09-935-625-598
Sequence 188190, A	C 326	16	0.8	1055	5	US-09-935-625-4320
Sequence 231937, A	C 327	16	0.8	1055	5	US-09-935-625-16506
Sequence 126527, A	C 328	16	0.8	1059	6	US-10-027-632-30791
Sequence 126528, A	C 329	16	0.8	1059	6	US-10-027-632-30792
Sequence 44023, A	C 330	16	0.8	1070	6	US-10-116-948-7
Sequence 101477, A	C 331	16	0.8	1095	7	US-60-360-039-26628
Sequence 22812, A	C 332	16	0.8	1120	6	US-10-027-632-263041
Sequence 2077, A	C 333	16	0.8	1120	6	US-10-027-632-263042
Sequence 2077, A	C 334	16	0.8	1140	7	US-60-360-039-42145
Sequence 2077, A	C 335	16	0.8	1161	7	US-60-360-039-43390
Sequence 248483, A	C 336	16	0.8	1192	7	US-60-360-039-40434
Sequence 204053, A	C 337	16	0.8	1198	6	US-10-027-632-296806
Sequence 204054, A	C 338	16	0.8	1215	7	US-60-360-039-39282
Sequence 7191, A	C 339	16	0.8	1215	7	US-60-360-039-39659
Sequence 224984, A	C 340	16	0.8	1215	7	US-60-360-039-40019
Sequence 228470, A	C 341	16	0.8	1245	6	US-10-027-632-122677
Sequence 228471, A	C 342	16	0.8	1245	6	US-10-027-632-122678
Sequence 228472, A	C 343	16	0.8	1245	6	US-10-027-632-122679
Sequence 228473, A	C 344	16	0.8	1245	6	US-10-027-632-109309
Sequence 228474, A	C 345	16	0.8	1275	6	US-10-027-632-208709
Sequence 228475, A	C 346	16	0.8	1292	5	US-09-935-625-9829
Sequence 151638, A	C 347	16	0.8	1292	5	US-09-935-625-22159
Sequence 151639, A	C 348	16	0.8	1292	5	US-09-935-625-22159
Sequence 151640, A	C 349	16	0.8	1308	5	US-09-935-625-9833
Sequence 151641, A	C 350	16	0.8	1308	5	US-09-935-625-22163
Sequence 32430, A	C 351	16	0.8	1308	5	US-60-360-039-26389
Sequence 25909, A	C 352	16	0.8	1314	5	US-09-935-625-8607
Sequence 164051, A	C 353	16	0.8	1353	7	US-60-360-039-43674
Sequence 164052, A	C 354	16	0.8	1361	7	US-60-360-039-47128
Sequence 151457, A	C 355	16	0.8	1371	7	US-60-360-039-25373
Sequence 803, A	C 356	16	0.8	1443	7	US-60-360-039-35658
Sequence 3179, A	C 357	16	0.8	1537	6	US-10-144-781-43
Sequence 28056, A	C 358	16	0.8	1538	5	US-09-935-625-7727
Sequence 98603, A	C 359	16	0.8	1570	5	US-09-919-002-3621
Sequence 98603, A	C 360	16	0.8	1584	5	US-09-935-625-11555
Sequence 119152, A	C 361	16	0.8	1584	5	US-09-935-625-19457
Sequence 119152, A	C 362	16	0.8	1584	5	US-09-935-625-27965
Sequence 29255, A	C 363	16	0.8	1595	6	US-10-105-299-2514
Sequence 12339, A	C 364	16	0.8	1604	6	US-10-027-632-249488
Sequence 162813, A	C 365	16	0.8	1605	7	US-60-360-039-26120
Sequence 174200, A	C 366	16	0.8	1632	7	US-60-360-039-35603
Sequence 174201, A	C 367	16	0.8	1646	1	PCT-US02-09188-1602
Sequence 156257, A	C 368	16	0.8	1646	1	PCT-US02-09370-1692
Sequence 7134, A	C 369	16	0.8	1646	6	US-10-105-299-7235
Sequence 156402, A	C 370	16	0.8	1652	1	PCT-US02-09188-335
Sequence 156403, A	C 371	16	0.8	1652	1	PCT-US02-09370-352
Sequence 156404, A	C 372	16	0.8	1652	6	US-10-105-299-466
Sequence 170387, A	C 373	16	0.8	1656	7	US-60-360-039-36017
Sequence 170388, A	C 374	16	0.8	1659	6	US-10-027-632-98305
Sequence 170389, A	C 375	16	0.8	1698	7	US-60-360-039-41349
Sequence 157911, A	C 376	16	0.8	1705	6	US-09-991-150-506
Sequence 157912, A	C 377	16	0.8	1705	6	US-10-121-062-393
Sequence 134116, A	C 378	16	0.8	1709	6	US-10-104-047-51
Sequence 120426, A	C 379	16	0.8	1727	7	US-60-360-039-46005
Sequence 120427, A	C 380	16	0.8	1743	6	US-10-006-063A-103
Sequence 167880, A	C 381	16	0.8	1743	6	US-10-006-117A-103
Sequence 167880, A	C 382	16	0.8	1743	6	US-10-006-130A-103
Sequence 120912, A	C 383	16	0.8	1743	6	US-10-006-172A-103
Sequence 46819, A	C 384	16	0.8	1743	6	US-10-006-768A-103
Sequence 37207, A	C 385	16	0.8	1743	6	US-10-017-522A-103
Sequence 4009, A	C 386	16	0.8	1743	6	US-10-017-610A-103
Sequence 153864, A	C 387	16	0.8	1743	6	US-10-006-041A-103
Sequence 153865, A	C 388	16	0.8	1743	6	US-10-006-818A-103
Sequence 153866, A	C 389	16	0.8	1743	6	US-10-012-121A-103
Sequence 110, A	C 390	16	0.8	1743	6	US-10-015-386A-103
Sequence 100933, A	C 391	16	0.8	1743	6	US-10-015-387A-103

Sequence 252940, A	319	16	0.8	991	6	US-10-027-632-252940
Sequence 31531, A	320	16	0.8	994	6	US-10-027-632-31531
Sequence 120200, A	C 321	16	0.8	1008	6	US-10-027-632-120200
Sequence 120201, A	C 322	16	0.8	1008	6	US-10-027-632-120201
Sequence 9749, A	C 323	16	0.8	1052	6	US-10-027-632-9749
Sequence 9750, A	C 324	16	0.8	1052	6	US-10-027-632-9750
Sequence 598, A	C 325	16	0.8	1055	5	US-09-935-625-598
Sequence 4320, A	C 326	16	0.8	1055	5	US-09-935-625-4320
Sequence 16506, A	C 327	16	0.8	1055	5	US-09-935-625-16506
Sequence 30791, A	C 328	16	0.8	1059	6	US-10-027-632-30791
Sequence 30792, A	C 329	16	0.8	1059	6	US-10-027-632-30792
Sequence 7, A	C 330	16	0.8	1070	6	US-10-116-948-7
Sequence 26628, A	C 331	16	0.8	1095	7	US-60-360-039-26628
Sequence 263041, A	C 332	16	0.8	1120	6	US-10-027-632-263041
Sequence 263042, A	C 333	16	0.8	1120	6	US-10-027-632-263042
Sequence 42145, A	C 334	16	0.8	1140	7	US

538	16	0.8	1743	6	US-10-063-723-75	Sequence 75, Appl	611	16	0.8	5320	5	US-09-442-384B-725	Sequence 725, App
539	16	0.8	1743	6	US-10-063-724-75	Sequence 75, Appl	612	16	0.8	6882	6	US-10-105-299-8509	Sequence 8509, Ap
540	16	0.8	1743	6	US-10-063-726-75	Sequence 75, Appl	613	16	0.8	8297	6	US-10-105-299-7938	Sequence 7938, Ap
541	16	0.8	1743	6	US-10-063-727-75	Sequence 75, Appl	614	16	0.8	8297	6	US-10-105-299-14263	Sequence 14263, A
542	16	0.8	1743	6	US-10-063-728-75	Sequence 75, Appl	615	16	0.8	10835	6	US-10-105-299-14988	Sequence 14988, A
543	16	0.8	1743	6	US-10-063-730-75	Sequence 75, Appl	616	16	0.8	10912	6	US-10-144-781-13	Sequence 13, Appl
544	16	0.8	1743	6	US-10-063-731-75	Sequence 75, Appl	617	16	0.8	11532	6	US-10-105-299-12161	Sequence 12161, A
545	16	0.8	1743	6	US-10-063-733-75	Sequence 75, Appl	618	16	0.8	11532	6	US-10-116-355-753	Sequence 753, App
546	16	0.8	1743	6	US-10-063-734-75	Sequence 75, Appl	619	16	0.8	13790	6	US-09-919-002-13473	Sequence 12473, A
547	16	0.8	1743	6	US-10-063-735-75	Sequence 75, Appl	620	16	0.8	13970	6	US-10-105-299-9643	Sequence 9643, Ap
548	16	0.8	1743	6	US-10-063-736-75	Sequence 75, Appl	621	16	0.8	14333	6	US-10-105-299-14987	Sequence 14987, A
549	16	0.8	1743	6	US-10-063-741-75	Sequence 75, Appl	622	16	0.8	15212	6	US-10-105-299-9506	Sequence 9506, Ap
550	16	0.8	1743	6	US-10-063-742-75	Sequence 75, Appl	623	16	0.8	17861	6	US-10-105-299-9930	Sequence 9930, Ap
551	16	0.8	1743	6	US-10-063-743-75	Sequence 75, Appl	624	16	0.8	21799	6	US-10-105-299-10314	Sequence 10314, A
552	16	0.8	1743	6	US-10-063-744-75	Sequence 75, Appl	625	16	0.8	21799	6	US-10-116-355-662	Sequence 662, App
553	16	0.8	1743	6	US-10-063-745-75	Sequence 75, Appl	626	16	0.8	25307	5	US-09-919-002-12002	Sequence 12002, A
554	16	0.8	1743	6	US-10-013-907A-103	Sequence 103, App	627	16	0.8	34505	6	US-10-117-983-479	Sequence 479, App
555	16	0.8	1743	6	US-10-015-391A-103	Sequence 103, App	628	16	0.8	35488	6	US-10-155-649-3	Sequence 3, Appl
556	16	0.8	1743	6	US-10-015-499A-103	Sequence 103, App	629	16	0.8	36620	1	PCT-US01-28861-25	Sequence 30, Appl
557	16	0.8	1743	6	US-10-017-867A-103	Sequence 103, App	630	16	0.8	37474	1	PCT-US01-28861-30	Sequence 25, Appl
558	16	0.8	1743	6	US-10-006-856A-103	Sequence 103, App	631	16	0.8	38519	1	PCT-US01-28861-38	Sequence 28, Appl
559	16	0.8	1757	1	PCT-US02-12851-2	Sequence 2, Appl	632	16	0.8	52156	5	US-09-578-519B-18	Sequence 18, Appl
560	16	0.8	1757	1	US-10-116-993-2	Sequence 2, Appl	633	16	0.8	160820	1	US-09-811-352B-1	Sequence 5, Appl
561	16	0.8	1758	6	US-10-144-850-32	Sequence 32, Appl	634	16	0.8	536165	5	US-09-939-964A-1	Sequence 1, Appl
562	16	0.8	1761	7	US-60-360-039-25611	Sequence 25611, A	635	16	0.8	1191139	6	US-10-140-924-1	Sequence 1, Appl
563	16	0.8	1793	5	US-09-919-002-12532	Sequence 12532, A	636	16	0.8	1191139	6	US-10-027-632-179264	Sequence 179264
564	16	0.8	1805	6	US-10-105-299-14869	Sequence 14869, A	637	16	0.8	1223197	6	US-10-067-514-1	Sequence 1, Appl
565	16	0.8	1836	7	US-60-360-039-25634	Sequence 25634, A	638	16	0.8	1691139	6	US-10-067-514-1	Sequence 1, Appl
566	16	0.8	1842	6	US-60-360-039-37001	Sequence 37001, A	639	16	0.8	1691139	6	US-10-067-514-1	Sequence 1, Appl
567	16	0.8	1844	6	US-10-027-633-97048	Sequence 97048, A	640	16	0.8	1691139	6	US-10-067-514-1	Sequence 1, Appl
568	16	0.8	1944	6	US-10-027-633-97949	Sequence 97949, A	641	15	0.8	27	6	US-10-027-632-174961	Sequence 174961
569	16	0.8	1944	6	US-10-027-633-98175	Sequence 98175, A	642	15	0.8	88	5	US-09-539-331D-38821	Sequence 38821, A
570	16	0.8	1944	6	US-10-027-633-98176	Sequence 98176, A	643	15	0.8	97	5	US-09-539-806B-311	Sequence 311, App
571	16	0.8	1944	6	US-10-027-633-98177	Sequence 98177, A	644	15	0.8	109	5	US-09-919-002-10017	Sequence 10017, A
572	16	0.8	1990	5	US-09-919-002-6691	Sequence 6691, Ap	645	15	0.8	136	5	US-09-540-210B-31013	Sequence 31013, A
573	16	0.8	2161	6	US-10-102-806-175	Sequence 175, App	646	15	0.8	155	5	US-09-539-331D-31261	Sequence 31261, A
574	16	0.8	2184	6	US-10-027-633-97571	Sequence 97571, A	647	15	0.8	169	6	US-10-141-324-1073	Sequence 1073, Ap
575	16	0.8	2188	6	US-10-099-926-1716	Sequence 1716, Ap	648	15	0.8	172	5	US-09-539-331D-8794	Sequence 8794, Ap
576	16	0.8	2197	6	US-10-116-355-606	Sequence 606, App	649	15	0.8	184	5	US-09-539-331D-36413	Sequence 36413, A
577	16	0.8	2268	5	US-10-099-926-1693	Sequence 1693, Ap	650	15	0.8	193	5	US-09-540-210B-22307	Sequence 22307, A
578	16	0.8	2269	5	US-09-954-531-955	Sequence 955, App	651	15	0.8	193	5	US-09-540-210B-28660	Sequence 28660, A
579	16	0.8	2271	5	US-09-935-625-21242	Sequence 21242, A	652	15	0.8	196	5	US-09-539-331D-26125	Sequence 26125, A
580	16	0.8	2329	5	US-09-882-227-619	Sequence 619, App	653	15	0.8	199	5	US-09-539-331D-12296	Sequence 12296, A
581	16	0.8	2352	6	US-10-027-633-262544	Sequence 262544, A	654	15	0.8	205	5	US-09-540-210B-23470	Sequence 23470, A
582	16	0.8	2391	7	US-60-360-039-24390	Sequence 24390, A	655	15	0.8	207	5	US-09-540-209B-16085	Sequence 16085, A
583	16	0.8	2423	6	US-10-104-047-1378	Sequence 1378, Ap	656	15	0.8	207	5	US-09-540-210B-16832	Sequence 16832, A
584	16	0.8	2479	6	US-10-027-633-103567	Sequence 103567, A	657	15	0.8	208	5	US-09-540-210B-16232	Sequence 16232, A
585	16	0.8	2485	6	US-10-027-633-111968	Sequence 111968, A	658	15	0.8	208	5	US-09-539-331D-36125	Sequence 36125, A
586	16	0.8	2498	6	US-10-027-633-102064	Sequence 102064, A	659	15	0.8	210	5	US-09-540-210B-30086	Sequence 30086, A
587	16	0.8	2498	6	US-10-027-633-102065	Sequence 102065, A	660	15	0.8	213	5	US-09-540-210B-14593	Sequence 14593, A
588	16	0.8	2498	6	US-10-027-633-102066	Sequence 102066, A	661	15	0.8	215	5	US-09-540-210B-575-81	Sequence 575, App
589	16	0.8	2530	1	PCT-US02-13644-5	Sequence 5, Appl	662	15	0.8	216	5	US-09-899-575-109	Sequence 81, Appl
590	16	0.8	2530	6	US-10-136-819-5	Sequence 5, Appl	663	15	0.8	217	5	US-09-540-210B-11350	Sequence 11350, A
591	16	0.8	2532	6	US-10-104-047-1069	Sequence 1069, Ap	664	15	0.8	218	5	US-09-540-210B-14969	Sequence 14969, A
592	16	0.8	2570	6	US-10-102-524-1764	Sequence 1764, Ap	665	15	0.8	218	6	US-10-146-502-1073	Sequence 1073, Ap
593	16	0.8	2760	5	US-09-935-625-16899	Sequence 16899, A	666	15	0.8	220	5	US-09-540-210B-6893	Sequence 6893, Ap
594	16	0.8	2918	5	US-09-919-002-13785	Sequence 3785, Ap	667	15	0.8	220	5	US-09-540-210B-30219	Sequence 30219, A
595	16	0.8	2962	5	US-09-919-002-1374	Sequence 1374, Ap	668	15	0.8	220	5	US-09-142-589-319	Sequence 319, App
596	16	0.8	3051	6	US-10-027-633-256743	Sequence 256743, A	669	15	0.8	226	5	US-09-973-234-17028	Sequence 17028, A
597	16	0.8	3051	6	US-10-027-633-256744	Sequence 256744, A	670	15	0.8	226	5	US-09-460-592B-1025	Sequence 1025, Ap
598	16	0.8	3183	7	US-60-360-039-35634	Sequence 35634, A	671	15	0.8	228	5	US-09-539-331D-4571	Sequence 4571, Ap
599	16	0.8	3416	5	US-09-930-213-293	Sequence 293, App	672	15	0.8	228	5	US-09-539-806B-556	Sequence 556, App
600	16	0.8	3497	6	US-10-027-633-114771	Sequence 114771, A	673	15	0.8	232	5	US-09-539-806C-754	Sequence 754, Ap
601	16	0.8	3608	6	US-10-121-062-433	Sequence 433, App	674	15	0.8	232	5	US-09-539-800C-3477	Sequence 3477, Ap
602	16	0.8	3692	6	US-10-106-698-824	Sequence 824, App	675	15	0.8	232	5	US-09-539-800C-8906	Sequence 8906, Ap
603	16	0.8	3754	6	US-10-104-047-530	Sequence 530, App	676	15	0.8	233	5	US-09-975-254-13113	Sequence 19113, A
604	16	0.8	3777	6	US-09-919-002-3576	Sequence 3576, Ap	677	15	0.8	233	5	US-09-539-331D-16546	Sequence 16546, A
605	16	0.8	4022	6	US-10-104-047-293	Sequence 293, App	678	15	0.8	233	5	US-09-539-331D-40504	Sequence 40504, A
606	16	0.8	4054	6	US-10-011-585A-32	Sequence 32, Appl	679	15	0.8	234	5	US-09-539-331D-17931	Sequence 17931, A
607	16	0.8	4145	5	US-09-053-375B-628	Sequence 628, App	680	15	0.8	235	5	US-09-539-800C-3992	Sequence 3992, Ap
608	16	0.8	4150	6	US-10-104-047-963	Sequence 963, App	681	15	0.8	236	5	US-09-539-800C-3992	Sequence 3992, Ap
609	16	0.8	4695	5	US-09-919-002-7377	Sequence 7377, Ap	682	15	0.8	236	6	US-10-027-633-289083	Sequence 289083, A
610	16	0.8	5188	6	US-10-105-299-9507	Sequence 9507, Ap	683	15	0.8	236	6	US-10-027-633-289084	Sequence 289084, A

c 684	15	0.8	241	5	US-09-540-2108-26184	Sequence 26184, A	c 757	15	0.8	351	5	US-09-895-9134-279	Sequence 279, App
c 685	15	0.8	242	5	US-09-975-254-24704	Sequence 24704, A	c 758	15	0.8	355	5	US-09-721-544-17650	Sequence 17650, A
c 686	15	0.8	243	5	US-09-975-254-4356	Sequence 4356, Ap	c 759	15	0.8	356	6	US-10-102-524-1048	Sequence 1048, Ap
c 687	15	0.8	243	5	US-09-539-331D-23443	Sequence 23443, A	c 760	15	0.8	359	6	US-10-103-313-119	Sequence 149, App
c 688	15	0.8	244	5	US-09-975-254-17067	Sequence 17067, A	c 761	15	0.8	359	6	US-10-027-632-269643	Sequence 269643, A
c 689	15	0.8	244	5	US-09-975-254-30774	Sequence 30774, A	c 762	15	0.8	359	6	US-10-027-632-269644	Sequence 269644, A
c 690	15	0.8	245	5	US-09-539-800C-18974	Sequence 18974, A	c 763	15	0.8	359	6	US-10-027-632-269645	Sequence 269645, A
c 691	15	0.8	246	5	US-09-539-331D-30196	Sequence 30196, A	c 764	15	0.8	360	5	US-09-721-544-16530	Sequence 16530, A
c 692	15	0.8	247	5	US-09-975-254-7622	Sequence 7622, Ap	c 765	15	0.8	368	5	US-09-721-544-8351	Sequence 8351, Ap
c 693	15	0.8	248	5	US-09-975-254-11792	Sequence 11792, A	c 766	15	0.8	368	5	US-09-721-544-13779	Sequence 13779, A
c 694	15	0.8	249	5	US-09-540-2108-1970	Sequence 1970, Ap	c 767	15	0.8	368	5	US-09-919-002-7642	Sequence 7642, Ap
c 695	15	0.8	250	5	US-09-540-2108-4520	Sequence 4520, Ap	c 768	15	0.8	369	7	US-60-360-039-37174	Sequence 37174, A
c 696	15	0.8	251	5	US-09-539-331D-8680	Sequence 8680, Ap	c 769	15	0.8	370	5	US-09-721-544-10057	Sequence 10057, A
c 697	15	0.8	251	5	US-09-539-800C-2885	Sequence 2885, Ap	c 770	15	0.8	374	5	US-09-539-331D-6828	Sequence 6828, Ap
c 698	15	0.8	252	5	US-09-975-254-13665	Sequence 13665, A	c 771	15	0.8	379	5	US-09-991-936-1111	Sequence 1111, Ap
c 699	15	0.8	253	5	US-09-539-331D-19235	Sequence 19235, A	c 772	15	0.8	380	5	US-09-919-002-10515	Sequence 10515, A
c 700	15	0.8	255	5	US-09-975-254-19783	Sequence 19783, A	c 773	15	0.8	387	5	US-09-919-002-11228	Sequence 11228, A
c 701	15	0.8	256	5	US-09-975-254-5198	Sequence 5198, Ap	c 774	15	0.8	390	5	US-09-721-544-23477	Sequence 23477, A
c 702	15	0.8	256	5	US-09-975-254-12792	Sequence 12792, A	c 775	15	0.8	392	5	US-09-721-544-14570	Sequence 14570, A
c 703	15	0.8	257	5	US-09-540-2108-7130	Sequence 7130, Ap	c 776	15	0.8	393	5	US-09-540-2098-4763	Sequence 4763, Ap
c 704	15	0.8	258	5	US-09-975-254-16817	Sequence 16817, A	c 777	15	0.8	393	5	US-09-919-002-10269	Sequence 10269, A
c 705	15	0.8	258	5	US-09-975-254-19449	Sequence 19449, A	c 778	15	0.8	394	5	US-09-171-937C-47	Sequence 47, Appl
c 706	15	0.8	260	5	US-09-975-254-24904	Sequence 24904, A	c 779	15	0.8	395	5	US-09-689-909-699	Sequence 699, App
c 707	15	0.8	260	5	US-09-941-492-73	Sequence 73, Appl	c 780	15	0.8	395	5	US-09-919-002-10189	Sequence 10189, A
c 708	15	0.8	262	5	US-09-539-331D-31245	Sequence 31245, A	c 781	15	0.8	396	6	US-10-027-632-183302	Sequence 183302, A
c 709	15	0.8	263	5	US-09-975-254-26125	Sequence 26125, A	c 782	15	0.8	399	6	US-10-114-893-320	Sequence 320, App
c 710	15	0.8	263	5	US-09-539-331D-7671	Sequence 7671, Ap	c 783	15	0.8	401	7	US-60-317-240-9854	Sequence 9854, Ap
c 711	15	0.8	264	5	US-09-975-254-11024	Sequence 11024, A	c 784	15	0.8	403	5	US-09-539-800C-14291	Sequence 14291, A
c 712	15	0.8	264	5	US-09-975-254-11036	Sequence 11036, A	c 785	15	0.8	403	6	US-10-113-872-1201	Sequence 1201, Ap
c 713	15	0.8	265	5	US-09-540-2108-12835	Sequence 12835, A	c 786	15	0.8	411	1	PCT-US02-12378-136	Sequence 136, App
c 714	15	0.8	266	5	US-09-975-254-2294	Sequence 2294, Ap	c 787	15	0.8	411	1	US-10-124-805-136	Sequence 136, App
c 715	15	0.8	266	5	US-09-539-331D-14182	Sequence 14182, A	c 788	15	0.8	412	6	US-10-011-154-302	Sequence 302, App
c 716	15	0.8	266	5	US-09-539-800C-18576	Sequence 18576, A	c 789	15	0.8	414	6	US-10-007-9266-55	Sequence 55, Appl
c 717	15	0.8	266	5	US-09-539-800C-18621	Sequence 18621, A	c 790	15	0.8	415	1	PCT-US02-10421-2068	Sequence 2068, Ap
c 718	15	0.8	267	5	US-09-975-254-6533	Sequence 6533, Ap	c 791	15	0.8	415	6	US-10-112-699-2068	Sequence 2068, Ap
c 719	15	0.8	267	5	US-09-540-2098-3348	Sequence 3348, Ap	c 792	15	0.8	415	6	US-10-027-632-183225	Sequence 183225, A
c 720	15	0.8	269	5	US-09-539-331D-295	Sequence 295, App	c 793	15	0.8	417	6	US-10-011-154-186	Sequence 186, App
c 721	15	0.8	269	5	US-09-539-806B-3157	Sequence 3157, Ap	c 794	15	0.8	418	6	US-10-027-632-299100	Sequence 299100, A
c 722	15	0.8	270	5	US-09-975-254-21506	Sequence 21506, A	c 795	15	0.8	419	5	US-09-919-002-5504	Sequence 5504, Ap
c 723	15	0.8	270	5	US-09-539-331D-38553	Sequence 38553, A	c 796	15	0.8	419	6	US-10-027-632-70184	Sequence 70184, A
c 724	15	0.8	271	5	US-09-539-331D-802	Sequence 802, App	c 797	15	0.8	420	6	US-10-027-632-277917	Sequence 277917, A
c 725	15	0.8	271	5	US-09-539-331D-13423	Sequence 13423, A	c 798	15	0.8	420	6	US-10-027-632-277918	Sequence 277918, A
c 726	15	0.8	272	5	US-09-975-254-9085	Sequence 9085, Ap	c 799	15	0.8	422	6	US-10-146-502-1828	Sequence 1828, Ap
c 727	15	0.8	277	5	US-09-539-800C-9689	Sequence 9689, Ap	c 800	15	0.8	422	6	US-10-027-632-43145	Sequence 43145, A
c 728	15	0.8	280	5	US-09-540-2108-26634	Sequence 26634, A	c 801	15	0.8	426	6	US-10-027-632-86524	Sequence 86524, A
c 729	15	0.8	280	5	US-09-539-331D-19436	Sequence 19436, A	c 802	15	0.8	426	6	US-10-027-632-86525	Sequence 86525, A
c 730	15	0.8	282	5	US-09-540-2108-13071	Sequence 13071, A	c 803	15	0.8	426	6	US-10-027-632-86526	Sequence 86526, A
c 731	15	0.8	286	5	US-09-539-331D-3335	Sequence 3335, Ap	c 804	15	0.8	426	6	US-10-027-632-316406	Sequence 316406, A
c 732	15	0.8	287	5	US-09-540-2108-33574	Sequence 33574, A	c 805	15	0.8	426	6	US-10-027-632-316407	Sequence 316407, A
c 733	15	0.8	287	5	US-09-539-806B-1493	Sequence 1493, Ap	c 806	15	0.8	426	6	US-10-027-632-316408	Sequence 316408, A
c 734	15	0.8	288	5	US-09-539-800C-13243	Sequence 13243, A	c 807	15	0.8	430	5	US-09-721-544-21591	Sequence 21591, A
c 735	15	0.8	292	1	PCT-US02-09188-1659	Sequence 1659, Ap	c 808	15	0.8	430	6	US-10-011-154-4557	Sequence 4557, Ap
c 736	15	0.8	292	1	PCT-US02-09370-1762	Sequence 1762, Ap	c 809	15	0.8	433	6	US-10-027-632-183144	Sequence 183144, A
c 737	15	0.8	292	5	US-09-540-2108-10133	Sequence 10133, A	c 810	15	0.8	433	6	US-10-027-632-183145	Sequence 183145, A
c 738	15	0.8	292	6	US-10-105-299-7325	Sequence 7325, Ap	c 811	15	0.8	433	6	US-10-027-632-183146	Sequence 183146, A
c 739	15	0.8	293	6	US-09-539-331D-31221	Sequence 31221, A	c 812	15	0.8	436	5	US-09-991-936-1630	Sequence 1630, Ap
c 740	15	0.8	294	5	US-09-540-2108-20446	Sequence 20446, A	c 813	15	0.8	436	6	US-10-011-154-4221	Sequence 4221, Ap
c 741	15	0.8	295	6	US-10-143-775-89	Sequence 89, Appl	c 814	15	0.8	437	5	US-09-721-544-21942	Sequence 21942, Ap
c 742	15	0.8	300	5	US-09-540-2108-35625	Sequence 35625, A	c 815	15	0.8	438	5	US-09-539-800C-16509	Sequence 16509, A
c 743	15	0.8	300	5	US-09-539-331D-35308	Sequence 35308, A	c 816	15	0.8	440	7	US-60-360-039-30766	Sequence 30766, A
c 744	15	0.8	301	6	US-10-027-632-35386	Sequence 35386, A	c 817	15	0.8	443	5	US-09-721-544-21331	Sequence 21331, A
c 745	15	0.8	301	6	US-10-027-632-65285	Sequence 65285, A	c 818	15	0.8	443	6	US-10-027-632-268517	Sequence 268517, A
c 746	15	0.8	308	6	US-10-144-654-6	Sequence 6, Appl1	c 819	15	0.8	443	6	US-10-027-632-268518	Sequence 268518, A
c 747	15	0.8	309	5	US-09-899-575-94	Sequence 94, Appl1	c 820	15	0.8	448	6	US-10-027-632-65820	Sequence 65820, A
c 748	15	0.8	310	5	US-09-539-331D-30380	Sequence 30380, A	c 821	15	0.8	448	6	US-10-027-632-65821	Sequence 65821, A
c 749	15	0.8	311	5	US-09-919-002-9159	Sequence 9159, Ap	c 822	15	0.8	448	6	US-10-027-632-66538	Sequence 66538, A
c 750	15	0.8	312	5	US-09-540-2108-10631	Sequence 10631, A	c 823	15	0.8	448	6	US-10-027-632-66539	Sequence 66539, A
c 751	15	0.8	318	6	US-10-116-355-644	Sequence 644, App	c 824	15	0.8	448	6	US-10-027-632-67466	Sequence 67466, A
c 752	15	0.8	323	5	US-09-941-492-102	Sequence 102, App	c 825	15	0.8	448	6	US-10-027-632-67467	Sequence 67467, A
c 753	15	0.8	335	5	US-09-539-331D-16731	Sequence 16731, A	c 826	15	0.8	448	6	US-10-027-632-67876	Sequence 67876, A
c 754	15	0.8	343	5	US-09-539-800C-9775	Sequence 9775, Ap	c 827	15	0.8	448	6	US-10-027-632-67877	Sequence 67877, A
c 755	15	0.8	347	5	US-09-721-544-20142	Sequence 20142, A	c 828	15	0.8	448	6	US-10-027-632-183658	Sequence 183658, A
c 756	15	0.8	348	5	US-09-721-544-10674	Sequence 10674, A	c 829	15	0.8	450	6	US-10-027-632-183692	Sequence 183692, A

C 830	15	0.8	450	6	US-10-027-632-78693	Sequence 78693, A	C 903	15	0.8	507	7	US-60-377-240-6237	Sequence 6237, Ap
C 831	15	0.8	450	6	US-10-027-632-314577	Sequence 314577, A	C 904	15	0.8	508	5	US-09-721-544-4927	Sequence 4927, Ap
C 832	15	0.8	450	6	US-10-027-632-314578	Sequence 314578, A	C 905	15	0.8	508	5	US-10-027-632-96959	Sequence 96959, A
C 833	15	0.8	452	6	US-10-011-154-1036	Sequence 1036, Ap	C 906	15	0.8	508	6	US-10-027-632-320980	Sequence 320980, A
C 834	15	0.8	453	5	US-09-620-393B-6564	Sequence 6564, Ap	C 907	15	0.8	510	6	US-10-027-632-78301	Sequence 78301, A
C 835	15	0.8	453	5	US-09-721-544-2193	Sequence 2193, A	C 908	15	0.8	510	6	US-10-027-632-78300	Sequence 78300, A
C 836	15	0.8	454	6	US-10-027-632-213709	Sequence 213709, A	C 909	15	0.8	510	6	US-10-027-632-78302	Sequence 78302, A
C 837	15	0.8	455	6	US-10-027-632-35196	Sequence 35196, A	C 910	15	0.8	510	6	US-10-027-632-79199	Sequence 79199, A
C 838	15	0.8	455	6	US-10-027-632-195204	Sequence 195204, A	C 911	15	0.8	510	6	US-10-027-632-79200	Sequence 79200, A
C 839	15	0.8	455	6	US-10-027-632-20893	Sequence 20893, A	C 912	15	0.8	510	6	US-10-027-632-79201	Sequence 79201, A
C 840	15	0.8	457	5	US-09-732-560-1	Sequence 1, Appl1	C 913	15	0.8	510	6	US-10-027-632-87203	Sequence 87203, A
C 841	15	0.8	458	1	PCT-US01-44838-1209	Sequence 1209, Ap	C 914	15	0.8	510	6	US-10-027-632-87204	Sequence 87204, A
C 842	15	0.8	458	6	US-10-027-632-288734	Sequence 288734, A	C 915	15	0.8	510	6	US-10-027-632-300742	Sequence 300742, A
C 843	15	0.8	459	5	US-09-919-002-11340	Sequence 11340, A	C 916	15	0.8	510	6	US-10-027-632-300743	Sequence 300743, A
C 844	15	0.8	459	6	US-10-027-632-85064	Sequence 85064, A	C 917	15	0.8	510	6	US-10-027-632-306912	Sequence 306912, A
C 845	15	0.8	460	5	US-09-919-002-6316	Sequence 6316, Ap	C 918	15	0.8	510	6	US-10-027-632-316606	Sequence 316606, A
C 846	15	0.8	460	6	US-10-027-632-81367	Sequence 81367, A	C 919	15	0.8	510	6	US-10-027-632-316607	Sequence 316607, A
C 847	15	0.8	462	6	US-10-027-632-209561	Sequence 209561, A	C 920	15	0.8	511	6	US-10-027-632-92073	Sequence 92073, A
C 848	15	0.8	466	7	US-60-377-240-7366	Sequence 7366, Ap	C 921	15	0.8	511	6	US-10-027-632-317886	Sequence 317886, A
C 849	15	0.8	468	5	US-09-689-909-1136	Sequence 1136, Ap	C 922	15	0.8	512	5	US-09-919-002-77209	Sequence 77209, Ap
C 850	15	0.8	468	6	US-10-027-632-194195	Sequence 194195, A	C 923	15	0.8	512	6	US-10-027-632-180687	Sequence 180687, A
C 851	15	0.8	468	6	US-10-027-632-194196	Sequence 194196, A	C 924	15	0.8	512	6	US-10-027-632-180688	Sequence 180688, A
C 852	15	0.8	470	6	US-10-027-632-28695	Sequence 28695, A	C 925	15	0.8	513	5	US-09-620-393B-7126	Sequence 7126, Ap
C 853	15	0.8	471	5	US-09-732-560-100	Sequence 100, Ap	C 926	15	0.8	514	5	US-09-539-800C-14552	Sequence 14552, A
C 854	15	0.8	471	5	US-09-721-544-11919	Sequence 11919, A	C 927	15	0.8	514	6	US-10-027-632-72540	Sequence 72540, A
C 855	15	0.8	473	6	US-10-027-632-273009	Sequence 273009, A	C 928	15	0.8	514	6	US-10-027-632-132275	Sequence 132275, A
C 856	15	0.8	474	6	US-10-027-632-70510	Sequence 70510, A	C 929	15	0.8	516	6	US-10-027-632-132276	Sequence 132276, A
C 857	15	0.8	476	6	US-10-027-632-181386	Sequence 181386, A	C 930	15	0.8	516	6	US-10-027-632-300803	Sequence 300803, A
C 858	15	0.8	476	6	US-10-027-632-181387	Sequence 181387, A	C 931	15	0.8	517	6	PCT-US02-10421-1219	Sequence 1219, Ap
C 859	15	0.8	476	6	US-10-027-632-181388	Sequence 181388, A	C 932	15	0.8	517	6	US-10-112-699-1219	Sequence 1219, Ap
C 860	15	0.8	477	6	US-10-027-632-34398	Sequence 34398, A	C 933	15	0.8	517	7	US-60-360-033-28798	Sequence 28798, A
C 861	15	0.8	477	6	US-10-027-632-34399	Sequence 34399, A	C 934	15	0.8	518	6	US-10-027-632-188217	Sequence 188217, A
C 862	15	0.8	479	6	US-10-027-632-64390	Sequence 64390, A	C 935	15	0.8	519	6	US-10-027-632-67043	Sequence 67043, A
C 863	15	0.8	480	5	US-09-721-544-11953	Sequence 11953, A	C 936	15	0.8	519	6	US-10-027-632-67044	Sequence 67044, A
C 864	15	0.8	481	5	US-09-721-544-7485	Sequence 7485, Ap	C 937	15	0.8	519	6	US-10-027-632-300803	Sequence 300803, A
C 865	15	0.8	482	7	US-60-377-240-5405	Sequence 5405, Ap	C 938	15	0.8	519	6	US-10-027-632-300804	Sequence 300804, A
C 866	15	0.8	483	5	US-09-539-331D-31802	Sequence 31802, A	C 939	15	0.8	520	5	US-09-620-393B-3374	Sequence 3374, A
C 867	15	0.8	483	5	US-09-721-544-7383	Sequence 7383, Ap	C 940	15	0.8	520	6	US-10-027-632-127843	Sequence 127843, A
C 868	15	0.8	483	6	US-10-027-632-83375	Sequence 83375, A	C 941	15	0.8	520	6	US-10-027-632-150888	Sequence 150888, A
C 869	15	0.8	485	5	US-09-721-544-3482	Sequence 3482, Ap	C 942	15	0.8	520	6	US-10-027-632-64867	Sequence 64867, A
C 870	15	0.8	485	6	US-10-027-632-42580	Sequence 42580, A	C 943	15	0.8	520	6	US-10-027-632-300594	Sequence 300594, A
C 871	15	0.8	489	5	US-09-919-002-458	Sequence 458, Ap	C 944	15	0.8	520	6	US-10-027-632-300594	Sequence 300594, A
C 872	15	0.8	489	7	US-60-377-240-4955	Sequence 4955, Ap	C 945	15	0.8	522	6	US-10-027-632-306895	Sequence 306895, A
C 873	15	0.8	493	6	US-10-027-632-274869	Sequence 274869, A	C 946	15	0.8	522	6	US-09-919-002-5108	Sequence 5108, Ap
C 874	15	0.8	493	6	US-10-027-632-321124	Sequence 321124, A	C 947	15	0.8	522	6	US-10-099-928-6659	Sequence 96364, A
C 875	15	0.8	493	6	US-10-027-632-321126	Sequence 321126, A	C 948	15	0.8	522	6	US-10-027-632-96365	Sequence 96365, A
C 876	15	0.8	496	5	US-09-721-544-250	Sequence 250, Ap	C 949	15	0.8	522	6	US-10-027-632-127841	Sequence 127841, A
C 877	15	0.8	496	6	US-10-027-632-246996	Sequence 246996, A	C 950	15	0.8	522	6	US-10-027-632-127843	Sequence 127843, A
C 878	15	0.8	497	6	US-09-539-331D-35400	Sequence 35400, A	C 951	15	0.8	522	6	US-10-027-632-127844	Sequence 127844, A
C 879	15	0.8	497	6	US-10-027-632-35041	Sequence 35041, A	C 952	15	0.8	522	6	US-10-027-632-270671	Sequence 270671, A
C 880	15	0.8	497	6	US-10-027-632-301352	Sequence 301352, A	C 953	15	0.8	522	6	US-10-027-632-306195	Sequence 306195, A
C 881	15	0.8	498	1	PCT-US02-12612-670	Sequence 670, Ap	C 954	15	0.8	522	6	US-10-027-632-306196	Sequence 306196, A
C 882	15	0.8	498	5	US-09-991-936-187	Sequence 187, Ap	C 955	15	0.8	523	6	US-10-027-632-81743	Sequence 81743, A
C 883	15	0.8	498	6	US-10-125-968-670	Sequence 670, Ap	C 956	15	0.8	523	6	US-10-027-632-81743	Sequence 81743, A
C 884	15	0.8	499	7	US-60-360-033-29432	Sequence 29432, A	C 957	15	0.8	523	6	US-10-027-632-93575	Sequence 93575, A
C 885	15	0.8	500	5	US-09-991-936-1767	Sequence 1767, Ap	C 958	15	0.8	523	6	US-10-027-632-307795	Sequence 307795, A
C 886	15	0.8	500	5	US-09-919-002-4908	Sequence 4908, Ap	C 959	15	0.8	524	5	US-09-539-800C-8711	Sequence 8711, Ap
C 887	15	0.8	500	6	US-10-027-632-71118	Sequence 71118, A	C 960	15	0.8	524	6	US-10-027-632-3562	Sequence 3562, Ap
C 888	15	0.8	500	6	US-10-027-632-245162	Sequence 245162, A	C 961	15	0.8	525	7	US-60-377-240-1714	Sequence 1714, Ap
C 889	15	0.8	500	6	US-10-027-632-245163	Sequence 245163, A	C 962	15	0.8	526	6	US-10-027-632-40630	Sequence 40630, A
C 890	15	0.8	500	6	US-10-027-632-287952	Sequence 287952, A	C 963	15	0.8	526	6	US-10-027-632-40631	Sequence 40631, A
C 891	15	0.8	500	6	US-10-027-632-287953	Sequence 287953, A	C 964	15	0.8	526	6	US-10-027-632-40632	Sequence 40632, A
C 892	15	0.8	502	6	US-10-027-632-223230	Sequence 223230, A	C 965	15	0.8	529	6	US-09-539-331D-34371	Sequence 34371, A
C 893	15	0.8	502	6	US-10-027-632-223231	Sequence 223231, A	C 966	15	0.8	529	6	US-10-027-632-37512	Sequence 37512, A
C 894	15	0.8	503	6	US-10-027-632-6807	Sequence 6807, Ap	C 967	15	0.8	529	6	US-10-027-632-37513	Sequence 37513, A
C 895	15	0.8	504	6	US-10-027-632-43952	Sequence 43952, A	C 968	15	0.8	529	6	US-10-027-632-334266	Sequence 334266, A
C 896	15	0.8	504	6	US-10-027-632-75620	Sequence 75620, A	C 969	15	0.8	529	6	US-10-027-632-334267	Sequence 334267, A
C 897	15	0.8	505	5	US-09-721-544-4928	Sequence 4928, Ap	C 970	15	0.8	530	5	US-09-539-331D-35157	Sequence 35157, A
C 898	15	0.8	505	6	US-10-027-632-290728	Sequence 290728, A	C 971	15	0.8	531	6	US-10-027-632-124282	Sequence 124282, A
C 899	15	0.8	505	6	US-10-027-632-290729	Sequence 290729, A	C 972	15	0.8	531	6	US-10-027-632-1344628	Sequence 1344628, A
C 900	15	0.8	505	6	US-10-027-632-290730	Sequence 290730, A	C 973	15	0.8	534	5	US-09-919-002-5585	Sequence 5585, Ap
C 901	15	0.8	506	5	US-09-539-331D-33413	Sequence 33413, A	C 974	15	0.8	537	5	US-09-540-208B-3138	Sequence 3138, Ap
C 902	15	0.8	507	6	US-10-027-632-300634	Sequence 300634, A	C 975	15	0.8	537	6	US-10-027-632-4508	Sequence 4508, Ap

c 976 15 0.8 537 6 US-10-027-632-4909 Sequence 4909, Ap
977 15 0.8 537 7 US-60-360-039-42617 Sequence 42617, A
978 15 0.8 539 5 US-09-721-544-5625 Sequence 5625, Ap
979 15 0.8 540 6 US-10-027-632-54028 Sequence 54028, A
980 15 0.8 540 6 US-10-027-632-289703 Sequence 289703, A
981 15 0.8 540 6 US-10-027-632-314558 Sequence 314558, A
982 15 0.8 540 6 US-10-027-632-321910 Sequence 321910, A
983 15 0.8 541 6 US-10-027-632-26193 Sequence 26193, A
984 15 0.8 541 6 US-10-027-632-109123 Sequence 109123, A
985 15 0.8 541 6 US-10-027-632-321125 Sequence 321125, A
986 15 0.8 542 6 US-10-027-632-203476 Sequence 203476, A
987 15 0.8 542 6 US-10-027-632-287844 Sequence 287844, A
988 15 0.8 543 6 US-10-027-632-285403 Sequence 285403, A
989 15 0.8 544 5 US-09-721-544-5626 Sequence 5626, Ap
990 15 0.8 545 6 US-10-027-632-76960 Sequence 76960, A
991 15 0.8 545 6 US-10-027-632-185376 Sequence 185376, A
992 15 0.8 545 6 US-10-027-632-300241 Sequence 300241, A
993 15 0.8 545 6 US-10-144-850-189 Sequence 189, App
994 15 0.8 546 6 US-10-027-632-41529 Sequence 41529, A
995 15 0.8 546 6 US-10-027-632-66628 Sequence 66628, A
996 15 0.8 546 6 US-10-027-632-210228 Sequence 210228, A
997 15 0.8 546 6 US-10-027-632-247833 Sequence 247833, A
998 15 0.8 547 6 US-10-027-632-204285 Sequence 204285, A
999 15 0.8 547 6 US-10-027-632-204286 Sequence 204286, A
c1000 15 0.8 547 6 US-10-027-632-204287 Sequence 204287, A

ALIGNMENTS

RESULT 1
US-08-961-083-1
Sequence 1, Application US/08961083
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961, 083
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1999 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-961-083-1

Query Match 100.0%; Score 1999; DB 4; Length 1999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAATTCAGACATTAATAATCACTCATCTGCTGATTTGGGTTTCGAACGGCCGCCA 60
1 TAAATTCAGACATTAATAATCACTCATCTGCTGATTTGGGTTTCGAACGGCCGCCA 60
Db 1 TAAATTCAGACATTAATAATCACTCATCTGCTGATTTGGGTTTCGAACGGCCGCCA 60
QY 61 TGGCCAAAGTAATGATTTCCACAGATTTGGTTAAGCAATGGTTTCATGCAAGCA 120
61 TGGCCAAAGTAATGATTTCCACAGATTTGGTTAAGCAATGGTTTCATGCAAGCA 120
Db 61 TGGCCAAAGTAATGATTTCCACAGATTTGGTTAAGCAATGGTTTCATGCAAGCA 120
QY 121 TCGCTTCTTGACACACAGGGGGATTGATACATCCGATTCCTGGAGCTTCTTGGCAA 180
121 TCGCTTCTTGACACACAGGGGGATTGATACATCCGATTCCTGGAGCTTCTTGGCAA 180
Db 121 TCGCTTCTTGACACACAGGGGGATTGATACATCCGATTCCTGGAGCTTCTTGGCAA 180
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATACCTCTACCCCAAGTGTATTAAGTTGAC 240
181 TCTGCAAGCAATTCCTCCCAAGGTGATACCTCTACCCCAAGTGTATTAAGTTGAC 240
Db 181 TCTGCAAGCAATTCCTCCCAAGGTGATACCTCTACCCCAAGTGTATTAAGTTGAC 240
QY 241 TTACTTTTCAACTTCGACTTCGACCCAGACACTATTTCTGTAAGGCTCAGAACTTGTT 300
241 TTACTTTTCAACTTCGACTTCGACCCAGACACTATTTCTGTAAGGCTCAGAACTTGTT 300
Db 241 TTACTTTTCAACTTCGACTTCGACCCAGACACTATTTCTGTAAGGCTCAGAACTTGTT 300
QY 301 AGCGATTCACTAGAACCAAAAGCAACCAAGCAAGAAATCTTACTATATAATTA 360
301 AGCGATTCACTAGAACCAAAAGCAACCAAGCAAGAAATCTTACTATATAATTA 360
Db 301 AGCGATTCACTAGAACCAAAAGCAACCAAGCAAGAAATCTTACTATATAATTA 360
QY 361 GGTCTACATGCTTAATGGAACATATGATGACAGACAGCTCAAAACTATATGTTAA 420
361 GGTCTACATGCTTAATGGAACATATGATGACAGACAGCTCAAAACTATATGTTAA 420
Db 361 GGTCTACATGCTTAATGGAACATATGATGACAGACAGCTCAAAACTATATGTTAA 420
QY 421 AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTCTGCTGGGAATGCTCAGGACC 480
421 AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTCTGCTGGGAATGCTCAGGACC 480
Db 421 AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTCTGCTGGGAATGCTCAGGACC 480
QY 481 AAACCAATATGACCCCTATTACATCCAGAGACGCCAAGCCGCGAACTTGCTT 540
481 AAACCAATATGACCCCTATTACATCCAGAGACGCCAAGCCGCGAACTTGCTT 540
Db 481 AAACCAATATGACCCCTATTACATCCAGAGACGCCAAGCCGCGAACTTGCTT 540
QY 541 ATCTGAATGAAAAATCAAGCTACATCTGCTGAAACATATATGAAAGCACTCATAC 600
541 ATCTGAATGAAAAATCAAGCTACATCTGCTGAAACATATATGAAAGCACTCATAC 600
Db 541 ATCTGAATGAAAAATCAAGCTACATCTGCTGAAACATATATGAAAGCACTCATAC 600
QY 601 ACCAATTACTGATGAGTACCAAGTCAATAGCAAGTAATTAACCTCTTACTATGA 660
601 ACCAATTACTGATGAGTACCAAGTCAATAGCAAGTAATTAACCTCTTACTATGA 660
Db 601 ACCAATTACTGATGAGTACCAAGTCAATAGCAAGTAATTAACCTCTTACTATGA 660
QY 661 TAAATTACCTGAAGGAATCATCAATCAAGTTGAAGAAAGCAAGCTATTAACCTACTAC 720
661 TAAATTACCTGAAGGAATCATCAATCAAGTTGAAGAAAGCAAGCTATTAACCTACTAC 720
Db 661 TAAATTACCTGAAGGAATCATCAATCAAGTTGAAGAAAGCAAGCTATTAACCTACTAC 720
QY 721 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAAAGCTCAAAACATGTGGGATAT 780
721 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAAAGCTCAAAACATGTGGGATAT 780
Db 721 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAAAGCTCAAAACATGTGGGATAT 780
QY 781 TTACAATACAGACGAATACGTTGCTATCCAGAGATGATTCGAATGCTCTTACTACAT 840
781 TTACAATACAGACGAATACGTTGCTATCCAGAGATGATTCGAATGCTCTTACTACAT 840
Db 781 TTACAATACAGACGAATACGTTGCTATCCAGAGATGATTCGAATGCTCTTACTACAT 840
QY 841 TGTGATGTTTCTAAGCGTAAGTCAATGGCCAGTGAAGACAGCCCATCAGTCAAGTAA 900
841 TGTGATGTTTCTAAGCGTAAGTCAATGGCCAGTGAAGACAGCCCATCAGTCAAGTAA 900
Db 841 TGTGATGTTTCTAAGCGTAAGTCAATGGCCAGTGAAGACAGCCCATCAGTCAAGTAA 900
QY 901 TGTTCCTTGGGAATTAACCAAGCAGTAAGAAACCCGAGCTGGGATCAACTATGAA 960
901 TGTTCCTTGGGAATTAACCAAGCAGTAAGAAACCCGAGCTGGGATCAACTATGAA 960
Db 901 TGTTCCTTGGGAATTAACCAAGCAGTAAGAAACCCGAGCTGGGATCAACTATGAA 960
QY 961 ACCGATACAGACTATGCTCTGCTGAGTACGGGTGTATACGATTCACCTGCTACTAT 1020
961 ACCGATACAGACTATGCTCTGCTGAGTACGGGTGTATACGATTCACCTGCTACTAT 1020
Db 961 ACCGATACAGACTATGCTCTGCTGAGTACGGGTGTATACGATTCACCTGCTACTAT 1020
QY 1021 CGTTCAAGATGAGCCCTATATACCTCGGAGCAAAATCTCTGTTATTAACCTGGATAG 1080


```

|||||
Db 1021 CGTTACAGATGACCCCTATTAACCTGAGACAAATACCTGTTATTAACGGGATAG 1080
QY 1081 GGGCTACTTTGGCAACATCACTTGCATACGCCCTGCAACAATCGCGAAGCTCCAGC 1140
Db 1081 GGGCTACTTTGGCAACATCACTTGCATACGCCCTGCAACAATCGCGAAGCTCCAGC 1140
QY 1141 CGTGAACACTCTTAACCAAGGTGAGTCAACCGCGCAAGACTTCTTAATGGTCTAGG 1200
Db 1141 CGTGAACACTCTTAACCAAGGTGAGTCAACCGCGCAAGACTTCTTAATGGTCTAGG 1200
QY 1201 AATGAGCTACCCCAAGTATTCATCTACTCAAAATGCAATTTCAAGTACACAACCGAATCAGA 1260
Db 1201 AATGAGCTACCCCAAGTATTCATCTACTCAAAATGCAATTTCAAGTACACAACCGAATCAGA 1260
QY 1261 CAAAAAATATGAGCAAGTAGTAAAAAGATGGTGTCTGTACGCTGCTTGGCAATGG 1320
Db 1261 CAAAAAATATGAGCAAGTAGTAAAAAGATGGTGTCTGTACGCTGCTTGGCAATGG 1320
QY 1321 TGGAACTTACTATTAACCAATGATATCCATTAAGTGGTCTTATGATGGGAGTGAAGA 1380
Db 1321 TGGAACTTACTATTAACCAATGATATCCATTAAGTGGTCTTATGATGGGAGTGAAGA 1380
QY 1381 AGAGTCTCTAATGTGGAAGCTGTGCAATGAAGAAACAGACAGCCTATATGATGACCGA 1440
Db 1381 AGAGTCTCTAATGTGGAAGCTGTGCAATGAAGAAACAGACAGCCTATATGATGACCGA 1440
QY 1441 CATGATGAAAAACAGTCTTACTTATGAACTGGACGAATGCTTACTTGGTCTGCC 1500
Db 1441 CATGATGAAAAACAGTCTTACTTATGAACTGGACGAATGCTTACTTGGTCTGCC 1500
QY 1501 TCAGGCTGTGTAAGAACAGAACTCTTACTATATACAGAGGAATGGAAGAACACATCA 1560
Db 1501 TCAGGCTGTGTAAGAACAGAACTCTTACTATATACAGAGGAATGGAAGAACACATCA 1560
QY 1561 GACCTCTCAATTTGTAGCACTGATGAACTATTTGCTGCTATATACGCTAATATTCAT 1620
Db 1561 GACCTCTCAATTTGTAGCACTGATGAACTATTTGCTGCTATATACGCTAATATTCAT 1620
QY 1621 GGGTATGAGACAGGCTATTTCAACCGTGTGACACCACTTGTAGCAATGGCTTACGGT 1680
Db 1621 GGGTATGAGACAGGCTATTTCAACCGTGTGACACCACTTGTAGCAATGGCTTACGGT 1680
QY 1681 CGGTGCAAGTTTACCGCTATGATGACCTGCTGTGGAAGAACATCCAGAGA 1740
Db 1681 CGGTGCAAGTTTACCGCTATGATGACCTGCTGTGGAAGAACATCCAGAGA 1740
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTAAATGGTCTCG 1800
Db 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTAAATGGTCTCG 1800
QY 1801 TTCTACGTGGAACCTCAACGCTCTCAACAACCCCATCACTCAAGTCAAGCTCAGC 1860
Db 1801 TTCTACGTGGAACCTCAACGCTCTCAACAACCCCATCACTCAAGTCAAGCTCAGC 1860
QY 1861 ATCAGATAGTCAACTTCAAGCTCTCAACGCTCTCAACGCTCTCAACGCTCTCAACGCT 1920
Db 1861 ATCAGATAGTCAACTTCAAGCTCTCAACGCTCTCAACGCTCTCAACGCTCTCAACGCT 1920
QY 1921 TACCAATCTCAACATTAATACGCAACATCAATATACACCCCTGATCAACAAATCAGAA 1980
Db 1921 TACCAATCTCAACATTAATACGCAACATCAATATACACCCCTGATCAACAAATCAGAA 1980
QY 1981 TCCTCAACGACACCA 1999
Db 1981 TCCTCAACGACACCA 1999

```

```

RESULT 2
US-08-961-083-227
; Sequence 227, Application US/08961083
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.

```

```

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 227:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 227:
US-08-961-083-227

```

```

Query Match 1.2%: Score 23; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAATCTACGACATATAATC 23
Db 11 TAAATCTACGACATATAATC 33

```

```

RESULT 3
US-09-540-209B-4827/c
; Sequence 4827, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 4827
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: B. fragilis
US-09-540-209B-4827

```

```

Query Match 1.1%: Score 21; DB 5; Length 1389;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 986 TTGAGTACGCTGTCTACAT 1006
Db 729 TTGAGTACGCTGTCTACAT 709

```



```
RESULT 4
US-09-721-544-12893/C
; Sequence 12893, Application US/09721544
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jensen, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Smith, Benjamin
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Fel
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
; FILE REFERENCE: 728CIP
; CURRENT APPLICATION NUMBER: US/09/721,544
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 09/515,128
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/034,341
; PRIOR FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12893
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(332)
; OTHER INFORMATION: n = A,T,C or G
US-09-721-544-12893

Query Match 1.0%; Score 20; DB 5; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1867 TAGTCACTTACAGCTCA 1886
Db 243 TAGTCACTTACAGCTCA 224

RESULT 5
US-10-027-632-240647
; Sequence 240647, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
```

```
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 240647
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-240647
```

```
Query Match 1.0%; Score 20; DB 6; Length 502;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 699 AACAGGCTATACCTACTC 718
Db 147 aaacagctataactactc 166
```

```
RESULT 6
US-10-027-632-240648
; Sequence 240648, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 240648
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-240648
```

```
Query Match 1.0%; Score 20; DB 6; Length 502;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 699 AACAGGCTATACCTACTC 718
Db 147 aaacagctataactactc 166
```



```
RESULT 7
US-10-027-632-214358
; Sequence 214358, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; SOFTWARE: FastSeq for Windows Version 4.0
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 214358
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-214358
```

```
Query Match 1.0%; Score 20; DB 6; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 699 AAACAGGCTATACCTACTC 718
DB 444 aaacaggtataactacc 463
```

```
RESULT 8
US-09-053-375B-301
; Sequence 301, Application US/09053375B
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; TITLE OF INVENTION: Nucleic Acid Arrays
; FILE REFERENCE: CION-006
; CURRENT APPLICATION NUMBER: US/09/053,375B
; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 1543
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 301
; LENGTH: 4508
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-053-375B-301
```

```
Query Match 1.0%; Score 20; DB 5; Length 4508;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 319 AAAGCAACCAAGCAAGAAA 338
DB 560 aaagcaaccacgaagaa 579
```

```
RESULT 9
US-09-620-393B-634/C
```

```
; Sequence 634, Application US/09620393B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1068P
; CURRENT APPLICATION NUMBER: US/09/620,393B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9948
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 634
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..427
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..427
; OTHER INFORMATION: Ceres Seq. ID 1377155
US-09-620-393B-634
```

```
Query Match 1.0%; Score 19; DB 5; Length 427;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1720 TGAAGGAAGCAATCCAGAA 1738
DB 78 TGAAGGAAGCAATCCAGAA 60
```

```
RESULT 10
US-09-620-393B-8776/C
; Sequence 8776, Application US/09620393B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1068P
; CURRENT APPLICATION NUMBER: US/09/620,393B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9948
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 8776
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..428
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..428
; OTHER INFORMATION: Ceres Seq. ID 1407382
US-09-620-393B-8776
```

```
Query Match 1.0%; Score 19; DB 5; Length 428;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1720 TGAAGGAAGCAATCCAGAA 1738
DB 78 TGAAGGAAGCAATCCAGAA 60
```

```
RESULT 11
US-10-027-632-98352
; Sequence 98352, Application US/10027632
```



```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98352
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-98352
```

```
Query Match 1.0%; Score 19; DB 6; Length 507;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 29 ATTGCTGACTGGTCTG 47
DB 319 attgcgacttggctctg 337
```

```
RESULT 12
US-10-027-632-250569/c
; Sequence 250569, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 250569
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-250569
```

```
Query Match 1.0%; Score 19; DB 6; Length 1079;
```

```
Best Local Similarity 100.0%; Pred. No. 7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 460 GGCTGGAATGCTCAGCA 478
DB 696 GGCTGGAATGCTCAGCA 678
```

```
RESULT 13
US-10-027-632-250570/c
; Sequence 250570, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 250570
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-250570
```

```
Query Match 1.0%; Score 19; DB 6; Length 1079;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 460 GGCTGGAATGCTCAGCA 478
DB 696 GGCTGGAATGCTCAGCA 678
```

```
RESULT 14
US-10-027-632-250571/c
; Sequence 250571, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
```


;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 250571
;; LENGTH: 1079
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-250571

Query Match 1.0%; Score 19; DB 6; Length 1079;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 460 GGCTGATGCTCAGCA 478
|||||
Db 696 GGCTGATGCTCAGCA 678

RESULT 15
US-60-360-039-23856
; Sequence 23856, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360.039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23856
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Xenorhabdus nematophilus
US-60-360-039-23856

Query Match 1.0%; Score 19; DB 7; Length 1299;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 99 CAATGTTCTATCGAGA 117
|||||
Db 857 caatglttctatcgaga 875

RESULT 16
US-60-360-039-24027
; Sequence 24027, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360.039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24027
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Xenorhabdus nematophilus
US-60-360-039-24027

Query Match 1.0%; Score 19; DB 7; Length 1299;

Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 99 CAATGTTCTATCGAGA 117
|||||
Db 857 caatglttctatcgaga 875

RESULT 17
US-60-360-039-42540
; Sequence 42540, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360.039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 42540
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Anabaena PCC7120
US-60-360-039-42540

Query Match 1.0%; Score 19; DB 7; Length 1920;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 184 GCAAGCAATTCCTCCAA 202
|||||
Db 923 gcaagcaatctcctccaa 941

RESULT 18
US-09-539-331D-25958
; Sequence 25958, Application US/09539331D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539.331D
; CURRENT FILING DATE: 2000-03-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO 25958
; LENGTH: 166
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00485767
US-09-539-331D-25958

Query Match 0.9%; Score 18; DB 5; Length 166;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1284 AAAAGATGCTGCTT 1301
|||||
Db 41 aaagatgctgctt 58


```
RESULT 19
US-09-539-800C-2484
; Sequence 2484, Application US/09539800C
; GENERAL INFORMATION:
; APPLICANT: Sellhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09/539,800C
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/521,383
; PRIOR FILING DATE: August 16, 1995
; PRIOR APPLICATION NUMBER: 08/271,217
; PRIOR FILING DATE: June 27, 1994
; PRIOR APPLICATION NUMBER: 08/334,881
; PRIOR FILING DATE: November 4, 1994
; PRIOR APPLICATION NUMBER: 08/943,978
; PRIOR FILING DATE: October 3, 1997
; PRIOR APPLICATION NUMBER: 60/028,732
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/943,979
; PRIOR FILING DATE: October 4, 1997
; PRIOR APPLICATION NUMBER: 60/027,782
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/993,774
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/034,975
; PRIOR FILING DATE: December 20, 1996
; PRIOR APPLICATION NUMBER: 09/250,003
; PRIOR FILING DATE: February 10, 1999
; PRIOR APPLICATION NUMBER: 60/074,364
; PRIOR FILING DATE: February 12, 1998
; PRIOR APPLICATION NUMBER: 09/452,747
; PRIOR FILING DATE: December 1, 1999
; PRIOR APPLICATION NUMBER: 60/111,910
; PRIOR FILING DATE: December 10, 1998
; NUMBER OF SEQ ID NOS: 19698
; SOFTWARE: PERL Program
; SEQ ID NO: 2484
; LENGTH: 229
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00209576
US-09-539-800C-2484

Query Match 0.9%; Score 18; DB 5; Length 229;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1284 AAAAGATGCTGCTGCTT 1301
DB 117 aaagatgctgctgctt 134
```

```
RESULT 20
US-09-540-210B-30401
; Sequence 30401, Application US/09540210B
; GENERAL INFORMATION:
; APPLICANT: Sellhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
; FILE REFERENCE: PD-1037 CIP
; CURRENT APPLICATION NUMBER: US/09/540,210B
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 08/972,899
; PRIOR FILING DATE: November 18, 1997
; PRIOR APPLICATION NUMBER: 08/395,244
; PRIOR FILING DATE: February 27, 1995
; PRIOR APPLICATION NUMBER: 08/722,922
; PRIOR FILING DATE: September 27, 1996
; PRIOR APPLICATION NUMBER: 60/005,526
; PRIOR FILING DATE: September 29, 1995
; PRIOR APPLICATION NUMBER: 08/824,029
; PRIOR FILING DATE: March 25, 1997
; PRIOR APPLICATION NUMBER: 60/014,010
; PRIOR FILING DATE: March 25, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/903,555
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/862,178
; PRIOR FILING DATE: May 22, 1997
; PRIOR APPLICATION NUMBER: 60/018,217
; PRIOR FILING DATE: May 23, 1996
; PRIOR APPLICATION NUMBER: 08/881,589
; PRIOR FILING DATE: June 24, 1997
; PRIOR APPLICATION NUMBER: 60/021,275
; PRIOR FILING DATE: June 25, 1996
; PRIOR APPLICATION NUMBER: 08/903,802
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/905,881
; PRIOR FILING DATE: August 1, 1997
; PRIOR APPLICATION NUMBER: 60/025,204
; PRIOR FILING DATE: August 1, 1996
; PRIOR APPLICATION NUMBER: 08/903,471
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025,478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025,217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937,142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026,598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960,746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030,144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755,524
; PRIOR FILING DATE: November 22, 1996
; PRIOR APPLICATION NUMBER: 60/007,495
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 09/021,031
; PRIOR FILING DATE: February 10, 1998
; PRIOR APPLICATION NUMBER: 60/039,325
; PRIOR FILING DATE: February 13, 1997
; PRIOR APPLICATION NUMBER: 09/035,172
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040,431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
```


Db 38 ATGACAGCTATTCTAA 21

RESULT 22

US-09-539-800C-15921

Sequence 15921, Application US/09539800C

GENERAL INFORMATION:

APPLICANT: Selihamer, Jeffrey J.

APPLICANT: Delegeane, Angelo M.

APPLICANT: Stuart, Susan G.

APPLICANT: Stuve, Laura L.

APPLICANT: Naughton, Rebecca E.

TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE

FILE REFERENCE: PD-1023 CTP

CURRENT APPLICATION NUMBER: US/09/539,800C

PRIOR FILING DATE: 2000-03-30

PRIOR APPLICATION NUMBER: 08/521,383

PRIOR FILING DATE: August 16, 1995

PRIOR APPLICATION NUMBER: 08/271,217

PRIOR FILING DATE: June 27, 1994

PRIOR APPLICATION NUMBER: 08/734,881

PRIOR FILING DATE: November 4, 1994

PRIOR APPLICATION NUMBER: 08/943,978

PRIOR FILING DATE: October 3, 1997

PRIOR APPLICATION NUMBER: 60/028,732

PRIOR FILING DATE: October 4, 1996

PRIOR APPLICATION NUMBER: 08/943,979

PRIOR FILING DATE: October 4, 1997

PRIOR APPLICATION NUMBER: 60/027,782

PRIOR FILING DATE: October 4, 1996

PRIOR APPLICATION NUMBER: 08/993,774

PRIOR FILING DATE: December 18, 1997

PRIOR APPLICATION NUMBER: 60/034,975

PRIOR FILING DATE: December 20, 1996

PRIOR APPLICATION NUMBER: 09/250,003

PRIOR FILING DATE: February 10, 1999

PRIOR APPLICATION NUMBER: 60/074,364

PRIOR FILING DATE: February 12, 1998

PRIOR APPLICATION NUMBER: 09/452,747

PRIOR FILING DATE: December 1, 1999

PRIOR APPLICATION NUMBER: 60/111,910

PRIOR FILING DATE: December 10, 1998

NUMBER OF SEQ ID NOS: 19698

SOFTWARE: PERL Program

SEQ ID NO: 15921

LENGTH: 235

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No: hu01253724

US-09-539-800C-15921

Query Match 0.9%; Score 18; DB 5; Length 250;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1284 AAAAGATGCTGCTGCTT 1301

Db 117 aaaagatgctgctgctt 134

RESULT 21

US-09-539-331D-33884/C

Sequence 33884, Application US/09539331D

GENERAL INFORMATION:

APPLICANT: Selihamer, Jeffrey J.

APPLICANT: Delegeane, Angelo M.

APPLICANT: Stuart, Susan G.

APPLICANT: Stuve, Laura L.

APPLICANT: Mullahy, Sara J.

APPLICANT: Naughton, Rebecca E.

TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE

FILE REFERENCE: PD-1022 CTP

CURRENT APPLICATION NUMBER: US/09/539,331D

PRIOR FILING DATE: 2000-03-30

Prior Application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 40961

SOFTWARE: PERL Program

SEQ ID NO: 33884

LENGTH: 250

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No: hu01166483

US-09-539-331D-33884

Query Match 0.9%; Score 18; DB 5; Length 250;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1627 ATGACAGCTATTCTAA 1644

|||||

Db 38 ATGACAGCTATTCTAA 21

RESULT 23

US-09-539-800C-16898

Sequence 16898, Application US/09539800C

GENERAL INFORMATION:

APPLICANT: Selihamer, Jeffrey J.

APPLICANT: Delegeane, Angelo M.

APPLICANT: Stuart, Susan G.

APPLICANT: Stuve, Laura L.

APPLICANT: Mullahy, Sara J.

Query Match 0.9%; Score 18; DB 5; Length 255;

Best Local Similarity 100.0%; Pred. No. 24;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1284 AAAAGATGCTGCTGCTT 1301

Db 201 aaaagatgctgctgctt 218

RESULT 22

US-09-539-800C-15921

Sequence 15921, Application US/09539800C

GENERAL INFORMATION:

APPLICANT: Selihamer, Jeffrey J.

APPLICANT: Delegeane, Angelo M.

APPLICANT: Stuart, Susan G.

APPLICANT: Stuve, Laura L.

APPLICANT: Naughton, Rebecca E.

TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE

FILE REFERENCE: PD-1023 CTP

CURRENT APPLICATION NUMBER: US/09/539,800C

PRIOR FILING DATE: 2000-03-30

PRIOR APPLICATION NUMBER: 08/521,383

PRIOR FILING DATE: August 16, 1995

PRIOR APPLICATION NUMBER: 08/271,217

PRIOR FILING DATE: June 27, 1994

PRIOR APPLICATION NUMBER: 08/734,881

PRIOR FILING DATE: November 4, 1994

PRIOR APPLICATION NUMBER: 08/943,978

PRIOR FILING DATE: October 3, 1997

PRIOR APPLICATION NUMBER: 60/028,732

PRIOR FILING DATE: October 4, 1996

PRIOR APPLICATION NUMBER: 08/943,979

PRIOR FILING DATE: October 4, 1997

PRIOR APPLICATION NUMBER: 60/027,782

PRIOR FILING DATE: October 4, 1996

PRIOR APPLICATION NUMBER: 08/993,774

PRIOR FILING DATE: December 18, 1997

PRIOR APPLICATION NUMBER: 60/034,975

PRIOR FILING DATE: December 20, 1996

PRIOR APPLICATION NUMBER: 09/250,003

PRIOR FILING DATE: February 10, 1999

PRIOR APPLICATION NUMBER: 60/074,364

PRIOR FILING DATE: February 12, 1998

PRIOR APPLICATION NUMBER: 09/452,747

PRIOR FILING DATE: December 1, 1999

PRIOR APPLICATION NUMBER: 60/111,910

PRIOR FILING DATE: December 10, 1998

NUMBER OF SEQ ID NOS: 19698

SOFTWARE: PERL Program

SEQ ID NO: 15921

LENGTH: 235

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No: hu01253724

US-09-539-800C-15921


```
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09/539,800C
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/521,383
; PRIOR FILING DATE: August 16, 1995
; PRIOR APPLICATION NUMBER: 08/271,217
; PRIOR FILING DATE: June 27, 1994
; PRIOR APPLICATION NUMBER: 08/334,881
; PRIOR FILING DATE: November 4, 1994
; PRIOR APPLICATION NUMBER: 08/943,978
; PRIOR FILING DATE: October 3, 1997
; PRIOR APPLICATION NUMBER: 60/028,732
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/943,979
; PRIOR FILING DATE: October 4, 1997
; PRIOR APPLICATION NUMBER: 60/027,782
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/993,774
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/034,975
; PRIOR FILING DATE: December 20, 1996
; PRIOR APPLICATION NUMBER: 09/250,003
; PRIOR FILING DATE: February 10, 1999
; PRIOR APPLICATION NUMBER: 60/074,364
; PRIOR FILING DATE: February 12, 1998
; PRIOR APPLICATION NUMBER: 09/452,747
; PRIOR FILING DATE: December 1, 1999
; PRIOR APPLICATION NUMBER: 60/111,910
; PRIOR FILING DATE: December 10, 1998
; NUMBER OF SEQ ID NOS: 19698
; SOFTWARE: PERL Program
; SEQ ID NO 16898
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu01110969
US-09-539-800C-16898

Query Match          0.9%; Score 18; DB 5; Length 264;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1284 AAAAGATGCTGCTGCTT 1301
Db 73 aaaagatgctgctgctt 90

RESULT 24
US-09-540-210B-24058
; Sequence 24058, Application US/09/540,210B
; GENERAL INFORMATION:
; APPLICANT: Sellhamey, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
; FILE REFERENCE: PD-1037 CIP
; CURRENT APPLICATION NUMBER: US/09/540,210B
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 08/972,899
; PRIOR FILING DATE: November 18, 1997
; PRIOR APPLICATION NUMBER: 08/395,244
; PRIOR FILING DATE: February 27, 1995
; PRIOR APPLICATION NUMBER: 08/722,922
; PRIOR FILING DATE: September 27, 1996
; PRIOR APPLICATION NUMBER: 60/005,526

; PRIOR FILING DATE: September 29, 1995
; PRIOR APPLICATION NUMBER: 08/824,029
; PRIOR FILING DATE: March 25, 1997
; PRIOR APPLICATION NUMBER: 60/014,010
; PRIOR FILING DATE: March 25, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/903,555
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/862,178
; PRIOR FILING DATE: May 22, 1997
; PRIOR APPLICATION NUMBER: 60/018,217
; PRIOR FILING DATE: May 23, 1996
; PRIOR APPLICATION NUMBER: 08/881,589
; PRIOR FILING DATE: June 24, 1997
; PRIOR APPLICATION NUMBER: 60/021,275
; PRIOR FILING DATE: June 25, 1996
; PRIOR APPLICATION NUMBER: 08/903,802
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/905,881
; PRIOR FILING DATE: August 1, 1997
; PRIOR APPLICATION NUMBER: 60/025,204
; PRIOR FILING DATE: August 1, 1996
; PRIOR APPLICATION NUMBER: 08/903,471
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025,478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025,217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937,142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026,598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960,746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030,144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755,524
; PRIOR FILING DATE: November 22, 1996
; PRIOR APPLICATION NUMBER: 60/007,495
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 09/021,031
; PRIOR FILING DATE: February 10, 1998
; PRIOR APPLICATION NUMBER: 60/039,325
; PRIOR FILING DATE: February 13, 1997
; PRIOR APPLICATION NUMBER: 09/035,172
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040,431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040,199
; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050,817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043,792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074,999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048,431
; PRIOR FILING DATE: May 29, 1997
```



```
; PRIOR APPLICATION NUMBER: 09/107,592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052,751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094,079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049,975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35654
; SOFTWARE: PERL Program
; SEQ ID NO 24058
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00169618
US-09-540-210B-24058
```

```
Query Match          0.9%; Score 18; DB 5; Length 268;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1585 TGAAGTATTGCTGCTA 1602
Db 74 tgacctattgtcgcta 91
```

```
RESULT 25
US-09-919-002-7641
; Sequence 7641, Application US/09919002
; GENERAL INFORMATION:
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/919,002
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: Application Number: 09/359,922
; PRIOR FILING DATE: FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: Application Number: US 09/034,341
; PRIOR FILING DATE: FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 13203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7641
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-002-7641
```

```
Query Match          0.9%; Score 18; DB 5; Length 277;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1783 ATTTAAATGCTGCTG 1800
Db 4 atttaaaatgctgctcg 21
```

```
RESULT 26
US-09-539-331D-21402
; Sequence 21402, Application US/09539331D
; GENERAL INFORMATION:
; APPLICANT: Selthamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
```

```
; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539,331D
; CURRENT FILING DATE: 2000-03-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO 21402
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00318340
; NAME/KEY: unsure
; LOCATION: 276, 281
; OTHER INFORMATION: a, t, c, g, or other
US-09-539-331D-21402
```

```
Query Match          0.9%; Score 18; DB 5; Length 285;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1284 AAAAGATGCTGCTGCTT 1301
Db 53 aaaagatgctgctgctt 70
```

```
RESULT 27
US-09-539-331D-34826
; Sequence 34826, Application US/09539331D
; GENERAL INFORMATION:
; APPLICANT: Selthamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539,331D
; CURRENT FILING DATE: 2000-03-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO 34826
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu01325546
US-09-539-331D-34826
```

```
Query Match          0.9%; Score 18; DB 5; Length 296;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1284 AAAAGATGCTGCTGCTT 1301
Db 197 aaaagatgctgctgctt 214
```

```
RESULT 28
US-10-027-632-94603
; Sequence 94603, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
```



```

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94603
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-94603
```

```

Query Match          0.9%; Score 18; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 TTCCTCCCAAGGTGATC 210
|||||
Db 436 ttccctcccaagtgatc 453
```

```

RESULT 29
US-10-027-632-94604
; Sequence 94604, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94604
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-94604
```

```

Query Match          0.9%; Score 18; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 TTCCTCCCAAGGTGATC 210
|||||
```

```

Db 436 ttccctcccaagtgatc 453
```

```

RESULT 30
US-10-027-632-305498
; Sequence 305498, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 305498
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305498
```

```

Query Match          0.9%; Score 18; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 193 TTCCTCCCAAGGTGATC 210
|||||
Db 436 ttccctcccaagtgatc 453
```

```

RESULT 31
US-10-027-632-305499
; Sequence 305499, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 305499
; LENGTH: 457
```



```
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305499

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 457;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 193 TTCCTCCAAGTGCATC 210
    |||
Db 436 ttcctccaagtgatc 453

RESULT 32
US-10-027-632-305500
; Sequence 305500, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 305500
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305500

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 457;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 193 TTCCTCCAAGTGCATC 210
    |||
Db 436 ttcctccaagtgatc 453

RESULT 33
US-10-027-632-286371/c
; Sequence 286371, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
```

```
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 286371
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-286371

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 478;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1530 ATACAGCAGGAATTG 1547
    |||
Db 201 ATACAGCAGGAATTG 184

RESULT 34
US-10-097-105-1069
; Sequence 1069, Application US/10097105
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Secrist, Heather
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.504C1
; CURRENT APPLICATION NUMBER: US/10/097,105
; PRIOR FILING DATE: 2002-03-13
; NUMBER OF SEQ ID NOS: 1562
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1069
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 318, 328, 347, 398, 400, 411, 423, 425, 429, 439, 443, 445,
; LOCATION: 458, 460, 478, 481, 498, 500, 502
; OTHER INFORMATION: n = A,T,C or G
US-10-097-105-1069

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 512;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1284 AAAAGATGCTGCTT 1301
    |||
Db 181 aaaagatgctgctt 198

RESULT 35
US-09-620-393B-6548
; Sequence 6548, Application US/09620393B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1068P
; CURRENT APPLICATION NUMBER: US/09/620,393B
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9948
```



```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6548
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..526
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..526
; OTHER INFORMATION: Ceres Seq. ID 1392936
; US-09-620-393B-6548

Query Match
Best Local Similarity 100.0%; Score 18; DB 5; Length 526;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1268 TATGAGCAAGTAGTGAA 1285
Db 248 tatgagcaagtagtgaa 265

RESULT 36
US-10-027-632-135305
; Sequence 135305, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIORITY FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 135305
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-135305

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 532;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1822 TCCACACAACCCCATC 1839
Db 74 tccacacaaccccatc 91

RESULT 37
US-10-027-632-81926/c
; Sequence 81926, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIORITY FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81926
; LENGTH: 581
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-81926

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 581;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 687 AAGTTGAGAGAAACAG 704
Db 95 AAGTTGAGAGAAACAG 78

RESULT 38
US-10-027-632-157285/c
; Sequence 157285, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIORITY FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 157285
; LENGTH: 850
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-157285

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 850;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


OY 966 TCACAGACTATGCTCTG 983
|||||
Db 635 TCACAGACTATGCTCTG 618

RESULT 39
US-10-106-698-791
; Sequence 791, Application US/10106698
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patentln Ver. 3.0
; SEQ ID NO 791
; LENGTH: 1229
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (19)..(19)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (305)..(305)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-791

Query Match 0.9%; Score 18; DB 6; Length 1229;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1627 ATGACAGGCTATCTAA 1644
|||||
Db 779 atgacagcgtatctctaa 796
RESULT 40
US-09-935-625-18639
; Sequence 18639, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 18639
; LENGTH: 1281
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: Misc.feature
; LOCATION: 1..1281
; OTHER INFORMATION: Ceres Seq. ID no. 1015072
US-09-935-625-18639

Query Match 0.9%; Score 18; DB 5; Length 1281;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1971 AAATCAGATCCTCAAC 1988
|||||

Db 676 aaatcagaatcctcaac 693

RESULT 41
US-09-935-625-27058
; Sequence 27058, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 27058
; LENGTH: 1281
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: Misc.feature
; LOCATION: 1..1281
; OTHER INFORMATION: Ceres Seq. ID no. 1015072
US-09-935-625-27058

Query Match 0.9%; Score 18; DB 5; Length 1281;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1971 AAATCAGATCCTCAAC 1988
|||||
Db 676 aaatcagaatcctcaac 693

RESULT 42
US-60-360-039-47071
; Sequence 47071, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 47071
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-60-360-039-47071

Query Match 0.9%; Score 18; DB 7; Length 1293;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 646 CCCTGCTACATGATTA 663
|||||
Db 399 ccctgcttacatgataa 416

RESULT 43
US-60-360-039-47214
; Sequence 47214, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)A
CURRENT APPLICATION NUMBER: US/60/360,039
CURRENT FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 47214
LENGTH: 1299
TYPE: DNA
ORGANISM: Escherichia coli
US-60-360-039-47214

Query Match 0.9%; Score 18; DB 7; Length 1299;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 101 ATCGTTCTATCGAAGAC 118
|||||

DB 856 atcgctctatcgagac 873

RESULT 44
US-09-919-002-13066/c
Sequence 13066, Application US/09919002
GENERAL INFORMATION:
APPLICANT: Leshowitz, Dena
APPLICANT: Liu, Jin
TITLE OF INVENTION: NOVEL CONFIGS OBTAINED FROM VARIOUS CDNA
LIBRARIES
FILE REFERENCE: 20411-752CON1
CURRENT APPLICATION NUMBER: US/09/919,002
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 09/359,922
PRIOR FILING DATE: FILING DATE: 1999-07-22
PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/034,341
PRIOR FILING DATE: FILING DATE: 1998-02-13
NUMBER OF SEQ ID NOS: 13203
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 13066
LENGTH: 1328
TYPE: DNA
ORGANISM: Homo sapiens
US-09-919-002-13066

Query Match 0.9%; Score 18; DB 5; Length 1328;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1284 AAAAGATGCTGCTT 1301
|||||

DB 525 AAAAGATGCTGCTT 508

RESULT 45
US-10-027-632-253001/c
Sequence 253001, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 253001
LENGTH: 1555
TYPE: DNA
ORGANISM: Human
US-10-027-632-253001

Query Match 0.9%; Score 18; DB 6; Length 1555;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 374 AATGGAACATATGGAATG 391
|||||

DB 704 AATGGAACATATGGAATG 687

RESULT 46
US-10-027-632-253002/c
Sequence 253002, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 253002
LENGTH: 1555
TYPE: DNA
ORGANISM: Human
US-10-027-632-253002

Query Match 0.9%; Score 18; DB 6; Length 1555;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 374 AATGGAACATATGGAATG 391
|||||

DB 704 AATGGAACATATGGAATG 687

RESULT 47
US-10-027-632-253003/c
Sequence 253003, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

;; TITLE OF INVENTION: Polymorphisms in the Human Genome
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 253003
;; LENGTH: 1555
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-253003

Query Match 0.9%: Score 18; DB 6; Length 1555;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 374 AATGGCACTATGATG 391
|||||
Db 704 AATGGCACTATGATG 687

RESULT 48
US-10-027-632-253004/C
;; Sequence 253004, Application US/10027632
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 253004
;; LENGTH: 1555
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-253004

Query Match 0.9%: Score 18; DB 6; Length 1555;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 374 AATGGCACTATGATG 391
|||||
Db 704 AATGGCACTATGATG 687

RESULT 49
US-60-360-039-24796/C
;; Sequence 24796, Application US/60360039
;; GENERAL INFORMATION:
;; APPLICANT: Cao, Yongwei
;; APPLICANT: Chen, Xianfeng
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; FILE REFERENCE: 38-10(52052)A
;; CURRENT APPLICATION NUMBER: US/60/360,039
;; CURRENT FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO: 24796
;; LENGTH: 1773
;; TYPE: DNA
;; ORGANISM: Methanobacterium thermoautotrophicum
US-60-360-039-24796

Query Match 0.9%: Score 18; DB 7; Length 1773;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1888 CTCACCACTCCACGAC 1905
|||||
Db 837 CTCACCACTCCACGAC 820

RESULT 50
US-10-115-123-53
;; Sequence 53, Application US/10115123
;; GENERAL INFORMATION:
;; APPLICANT: Ruben et al.
;; TITLE OF INVENTION: 94 Human Secreted Proteins
;; FILE REFERENCE: P2029630AIP12
;; CURRENT APPLICATION NUMBER: US/10/115,123
;; CURRENT FILING DATE: 2002-04-04
;; PRIOR APPLICATION NUMBER: PCT/US99/13418
;; PRIOR FILING DATE: 1999-06-15
;; PRIOR APPLICATION NUMBER: 60/089,507
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089,508
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089,509
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089,510
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/090,112
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090,113
;; PRIOR FILING DATE: 1998-06-22
;; NUMBER OF SEQ ID NOS: 532
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 53
;; LENGTH: 2288
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (940)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: SITE
;; LOCATION: (1279)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: SITE

Query Match	0.9%	Score 18	DB 6	Length 2288
Best Local Similarity	100.0%	Pred. No. 28		
Matches 18; Conservative	0	Mismatches	0	Indels 0; Gaps 0

	Query Match	0.9%	Score 18;	DB 6;	Length 2288;
	Best Local Similarity	100.0%;	Pred. No. 28;		
Matches	18; Conservative	0;	Mismatches	0;	Indels 0;
Gaps					0;
OY	1284 AAAAGATGCGTCTGCTT	1301			
Db	1734 aaaagatgcgtcgtctt	1751			

Search completed: June 13, 2002, 13:44:17
Job time: 7794 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:40:34 ; Search time 21.48 Seconds
(without alignments)
2979.306 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484

Sequence: 1 KIYDNKNQIADLGSESRVYN.....TQOSNTTTPQONQNPQAPQ 666

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3484	100.0	719	2	G95042
2	3475	99.7	719	2	S28038
3	3471	99.6	719	2	S28037
4	3466	99.5	719	2	A42893
5	3466	99.5	719	2	A97913
6	3246	93.2	719	2	S28034
7	3245	93.1	719	2	S28032
8	3243	93.1	719	2	S28031
9	3241	93.0	719	2	S28033
10	3192	91.6	719	2	S28035
11	2820	80.9	608	2	S28036
12	2698.5	77.5	637	2	B42893
13	1657.5	47.6	664	2	G86692
14	1217	34.9	827	2	AD1311
15	1201	34.5	826	2	AD1683
16	1074.5	30.8	886	2	F83862
17	999	28.7	914	2	I40529
18	914	26.2	188	2	S31952
19	817.5	23.5	727	2	B89923
20	813.5	23.3	716	2	S43693
21	802.5	23.0	809	2	F97183
22	739.5	21.2	714	2	AE1353
23	739	21.2	714	2	AE1353
24	724	20.8	714	2	AE1353
25	712	20.4	726	2	F70355
26	698.5	20.0	719	2	A83800
27	670.5	19.2	643	2	EB1396
28	656	18.8	966	2	EB4053
29	655	18.8	764	2	B97371

30	655	18.8	764	2	AB2589	penicillin-binding
31	649.5	18.6	743	2	D86888	penicillin-binding
32	642.5	18.4	822	2	F83016	penicillin-binding
33	640	18.4	798	2	H81040	penicillin-binding
34	637	18.3	731	2	D95235	penicillin-binding
35	636	18.3	731	2	E98099	penicillin-binding
36	627	18.0	659	2	E64594	penicillin-binding
37	625	17.9	660	2	F71917	penicillin-binding
38	618	17.7	851	2	AF0018	penicillin-binding
39	617	17.7	835	2	B82051	penicillin-binding
40	615	17.7	735	2	AI2174	penicillin-binding
41	614	17.6	873	2	B75514	penicillin-binding
42	613.5	17.6	687	2	D84126	penicillin-binding
43	608	17.5	718	2	AI3420	penicillin-binding
44	598.5	17.2	624	2	A55220	penicillin-binding
45	598	17.2	885	2	S76357	penicillin-binding

ALIGNMENTS

RESULT 1

G95042 penicillin-binding protein 1A [imported] - Streptococcus pneumoniae (strain TIGR4)

C/Species: Streptococcus pneumoniae

C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001

C/Accession: G95042

R/Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf

erson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris

A>Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000, MUID:21357209, PMID:11463916

A:Accession: G95042

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-719 <KUR>

A:Cross-references: GB:AE005672; PIDN:AAK74536.1; PID:gl4971838; GSPDB:GN00164; TIGR:

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP0369

C:Superfamily: penicillin-binding protein 1B

Query Match	Best Local Similarity	Score	DB 2:	Length	719:
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
OY 1	KIYDNKNQIADLGSESRVNAQNDIPTDLVKAIVSIEDHRRFPDRGIDTIRIIGAFARN 60				
DB 54	KIYDNKNQIADLGSESRVNAQNDIPTDLVKAIVSIEDHRRFPDRGIDTIRIIGAFARN 113				
OY 61	LOSNSLOGSTLTQOLIKLTFYFSTSDQTSRKAQEMLAIOEQKATKOEILTYINK 120				
DB 114	LOSNSLOGSTLTQOLIKLTFYFSTSDQTSRKAQEMLAIOEQKATKOEILTYINK 173				
OY 121	VYNSNGNGMGTAAQNTYGRKLNLSLPQALLAGMPAPQYDPYSHPEAAQDRNLVL 180				
DB 174	VYNSNGNGMGTAAQNTYGRKLNLSLPQALLAGMPAPQYDPYSHPEAAQDRNLVL 233				
OY 181	SEMKNGYISAEQYKAVNPITDGLQSLKASNPVPMQVYLVKEVINQVEETGYNLLT 240				
DB 234	SEMKNGYISAEQYKAVNPITDGLQSLKASNPVPMQVYLVKEVINQVEETGYNLLT 293				
OY 241	TGMVYTNVDEQAKHLMIDYNTDEYVAYPPDELQVASTIVDSNGKVIQAQARHSSN 300				
DB 294	TGMVYTNVDEQAKHLMIDYNTDEYVAYPPDELQVASTIVDSNGKVIQAQARHSSN 353				
OY 301	VSFGINQAVETNRMGSTMKRPTDYAPALEGVYDSTATYVHDEPYNPGTNTPYVYNNDR 360				
DB 354	VSFGINQAVETNRMGSTMKRPTDYAPALEGVYDSTATYVHDEPYNPGTNTPYVYNNDR 413				

QY 361 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFPLNGLGIDYPSIHSNAISSNTTESD 420
|||||
Db 414 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFPLNGLGIDYPSIHSNAISSNTTESD 473
QY 421 KRYGASSEMAAAYAAAFANGGTYKPMYIHKYVFSGSGSEKESNVGTRAMKETTAAMMD 480
|||||
Db 474 KRYGASSEMAAAYAAAFANGGTYKPMYIHKYVFSGSGSEKESNVGTRAMKETTAAMMD 533
QY 481 MKKTVLTGGRNAYLAMLPOAGKTGTSNYTDEBEIENHIKTSOFVAPDELFGAGYTRKYSM 540
|||||
Db 534 MKKTVLTGGRNAYLAMLPOAGKTGTSNYTDEBEIENHIKTSOFVAPDELFGAGYTRKYSM 593
QY 541 AVWTGYSNRLTPLYGNGLTVAAKVYRSAMTYLSEGSNPEDWNIPEGLYRNGEFVFNKGAR 600
|||||
Db 594 AVWTGYSNRLTPLYGNGLTVAAKVYRSAMTYLSEGSNPEDWNIPEGLYRNGEFVFNKGAR 653
QY 601 STWNSPAPQOPPTSESSSSSDSSTPSTNNSTTTNPNNTTQOOSNTTTPDOONON 660
|||||
Db 654 STWNSPAPQOPPTSESSSSSDSSTPSTNNSTTTNPNNTTQOOSNTTTPDOONON 713
QY 661 POPAP 666
|||||
Db 714 POPAP 719

RESULT 2
S28038
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 45607) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28038
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of pen
A:Reference number: S28031; MUID:93010977
A:Accession: S28038
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67873; NID:q47419; PIDN:CAA48073.1; PID:q47420
C:Superfamily: penicillin-binding protein 1B

Query Match 99.7%; Score 3475; DB 2; Length 719;
Best Local Similarity 99.7%; Pred. No. 1.5e-193;
Matches 664; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 KIYDNKNQLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 60
|||||
Db 54 KIYDNKNQLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 113
QY 61 LOSNSLOGGSTLTQOLIKLTFFSTSDQTSRKAQEAHLAQLQEKATKQELTYIYINK 120
|||||
Db 114 LOSNSLOGGSTLTQOLIKLTFFSTSDQTSRKAQEAHLAQLQEKATKQELTYIYINK 173
QY 121 VYMSNGNYGMQTAQNYKYKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRNNLV 180
|||||
Db 174 VYMSNGNYGMQTAQNYKYKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRNNLV 233
QY 181 SEMKNQGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLT 240
|||||
Db 234 SEMKNQGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLT 293
QY 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVVAYPDDELQVASTIVDSNGKVIAQLGARHOSN 300
|||||
Db 294 TGMDEVYTNVDOEAQKHLMDIYNTDEVVAYPDDELQVASTIVDSNGKVIAQLGARHOSN 353
QY 301 VSEFGINQAVETNRDMSGTMKPTIDYAPALEYGYVDSSTATIVHDEPNYNGTNPVYNNMR 360
|||||
Db 354 VSEFGINQAVETNRDMSGTMKPTIDYAPALEYGYVDSSTATIVHDEPNYNGTNPVYNNMR 413
QY 361 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFPLNGLGIDYPSIHSNAISSNTTESD 420
|||||

Db 414 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFPLNGLGIDYPSIHSNAISSNTTESD 473
QY 421 KRYGASSEMAAAYAAAFANGGTYKPMYIHKYVFSGSGSEKESNVGTRAMKETTAAMMD 480
|||||
Db 474 KRYGASSEMAAAYAAAFANGGTYKPMYIHKYVFSGSGSEKESNVGTRAMKETTAAMMD 533
QY 481 MKKTVLTGGRNAYLAMLPOAGKTGTSNYTDEBEIENHIKTSOFVAPDELFGAGYTRKYSM 540
|||||
Db 534 MKKTVLTGGRNAYLAMLPOAGKTGTSNYTDEBEIENHIKTSOFVAPDELFGAGYTRKYSM 593
QY 541 AVWTGYSNRLTPLYGNGLTVAAKVYRSAMTYLSEGSNPEDWNIPEGLYRNGEFVFNKGAR 600
|||||
Db 594 AVWTGYSNRLTPLYGNGLTVAAKVYRSAMTYLSEGSNPEDWNIPEGLYRNGEFVFNKGAR 653
QY 601 STWNSPAPQOPPTSESSSSSDSSTPSTNNSTTTNPNNTTQOOSNTTTPDOONON 660
|||||
Db 654 STWNSPAPQOPPTSESSSSSDSSTPSTNNSTTTNPNNTTQOOSNTTTPDOONON 713
QY 661 POPAP 666
|||||
Db 714 POPAP 719

RESULT 3
S28037
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 63915) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28037
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of
A:Reference number: S28031; MUID:93010977
A:Accession: S28037
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67872; NID:q47417; PIDN:CAA48072.1; PID:q47418
C:Superfamily: penicillin-binding protein 1B

Query Match 99.6%; Score 3471; DB 2; Length 719;
Best Local Similarity 99.4%; Pred. No. 2.6e-193;
Matches 662; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 KIYDNKNQLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 60
|||||
Db 54 KIYDNKNQLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 113
QY 61 LOSNSLOGGSTLTQOLIKLTFFSTSDQTSRKAQEAHLAQLQEKATKQELTYIYINK 120
|||||
Db 114 LOSNSLOGGSTLTQOLIKLTFFSTSDQTSRKAQEAHLAQLQEKATKQELTYIYINK 173
QY 121 VYMSNGNYGMQTAQNYKYKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRNNLV 180
|||||
Db 174 VYMSNGNYGMQTAQNYKYKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRNNLV 233
QY 181 SEMKNQGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLT 240
|||||
Db 234 SEMKNQGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLT 293
QY 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVVAYPDDELQVASTIVDSNGKVIAQLGARHOSN 300
|||||
Db 294 TGMDEVYTNVDOEAQKHLMDIYNTDEVVAYPDDELQVASTIVDSNGKVIAQLGARHOSN 353
QY 301 VSEFGINQAVETNRDMSGTMKPTIDYAPALEYGYVDSSTATIVHDEPNYNGTNPVYNNMR 360
|||||
Db 354 VSEFGINQAVETNRDMSGTMKPTIDYAPALEYGYVDSSTATIVHDEPNYNGTNPVYNNMR 413
QY 361 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFPLNGLGIDYPSIHSNAISSNTTESD 420
|||||
Db 414 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFPLNGLGIDYPSIHSNAISSNTTESD 473
|||||


```
QY 421 KKGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKFEFNVGTAKMETTAYAMTD 480
|||||
DB 474 KKGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKFEFNVGTAKMETTAYAMTD 533
QY 481 MKKTYLTGTGRNAYLAAPQAGKTGTSNTYDELENIHKTQSYAPADELFAGTTRKXSM 540
|||||
DB 534 MKKTYLSTGTGRNAYLAAPQAGKTGTSNTYDELENIHKTQSYAPADELFAGTTRKXSM 593
QY 541 AVMTGYSNRLLPVLVNGSLTVAAYKYSRSMWTYLSGSGNPEDMNIPGGLYRNGEFPFKNGAR 600
|||||
DB 594 AVMTGYSNRLLPVLVNGSLTVAAYKYSRSMWTYLSGSGNPEDMNIPGGLYRNGEFPFKNGAR 653
QY 601 STWNSPAPQPPSTRESSSSSDSSSTQSSSTPTSTNNSTTNPNNNTQOQSWTTPDOQON 660
|||||
DB 654 STWNSPAPQPPSTRESSSSSDSSSTQSSSTPTSTNNSTTNPNNNTQOQSWTTPDOQON 713
QY 661 POPAP 666
|||||
DB 714 POPAP 719
```

```
RESULT 4
A42893
penicillin-binding protein 1A - Streptococcus pneumoniae
C:Species: Streptococcus pneumoniae
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 21-Jan-2000
C:Accession: A42893
R:Martin, C.; Briese, T.; Hakenbeck, R.
J. Bacteriol. 174, 4517-4523, 1992
A:Title: Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus pneumoniae.
A:Reference number: A42893; MUID:92325042
A:Accession: A42893
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: GB:M90527; NID:g153766; PIDN:AAA26956.1; PID:g153768
C:Superfamily: penicillin-binding protein 1B
```

```
Query Match 99.5%; Score 3466; DB 2; Length 719;
Best Local Similarity 99.4%; Pred. No. 5.1e-193;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLKAYISIEDHFPDRGIDTRIIGAFLRN 60
|||||
DB 54 KIYDNKNQIADLGSERRVNAQANDIPTDLKAYISIEDHFPDRGIDTRIIGAFLRN 113
QY 61 LOSNSLOGGSLTQOLIKLTFYFSTSDQTSRKAQEWALAIQLEOKATKOEILTYINK 120
|||||
DB 114 LOSNSLOGGSLTQOLIKLTFYFSTSDQTSRKAQEWALAIQLEOKATKOEILTYINK 173
QY 121 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPQAPQYDPSHPEAQAORRNLYL 180
|||||
DB 174 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPQAPQYDPSHPEAQAORRNLYL 233
QY 181 SEMKNQGYISAEQYKAVNPITDGLQSLKASNPYAMNDVYLKEVINQVEEFGYNLLT 240
|||||
DB 234 SEMKNQGYISAEQYKAVNPITDGLQSLKASNPYAMNDVYLKEVINQVEEFGYNLLT 293
QY 241 TGMVYTNVDEQAKHLMIDYNTDEYVAYPDDELQVASTIVDVSNKGYIAQLGARHSSN 300
|||||
DB 294 TGMVYTNVDEQAKHLMIDYNTDEYVAYPDDELQVASTIVDVSNKGYIAQLGARHSSN 353
QY 301 VSEGINQAVETNRDQWGTMRPITDYAPALEGYVDSTATYHDEPYNPGTNTPYNNMR 360
|||||
DB 354 VSEGINQAVETNRDQWGTMRPITDYAPALEGYVDSTATYHDEPYNPGTNTPYNNMR 413
QY 361 GYFGNITLQYALQOOSRNPAVETLNKVGILNRAKTFNLGLIDYPSIHYSNAISSNTESD 420
|||||
DB 414 GYFGNITLQYALQOOSRNPAVETLNKVGILNRAKTFNLGLIDYPSIHYSNAISSNTESD 473
QY 421 KKGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKFEFNVGTAKMETTAYAMTD 480
```

```
DB 474 KKGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKFEFNVGTAKMETTAYAMTD 533
|||||
QY 481 MKKTYLTGTGRNAYLAAPQAGKTGTSNTYDELENIHKTQSYAPADELFAGTTRKXSM 540
|||||
DB 534 MKKTYLSTGTGRNAYLAAPQAGKTGTSNTYDELENIHKTQSYAPADELFAGTTRKXSM 593
QY 541 AVMTGYSNRLLPVLVNGSLTVAAYKYSRSMWTYLSGSGNPEDMNIPGGLYRNGEFPFKNGAR 600
|||||
DB 594 AVMTGYSNRLLPVLVNGSLTVAAYKYSRSMWTYLSGSGNPEDMNIPGGLYRNGEFPFKNGAR 653
QY 601 STWNSPAPQPPSTRESSSSSDSSSTQSSSTPTSTNNSTTNPNNNTQOQSWTTPDOQON 660
|||||
DB 654 STWNSPAPQPPSTRESSSSSDSSSTQSSSTPTSTNNSTTNPNNNTQOQSWTTPDOQON 713
QY 661 POPAP 666
|||||
DB 714 POPAP 719
```

```
RESULT 5
A97913
peptidoglycan glycosyltransferase (EC 2.4.1.129) [imported] - Streptococcus pneumonia
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: A97913
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAnthen, S.
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A97913
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-719 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAK9133.1; PID:g1545786; GSPDB:GN00174
C:Genetics:
A:Gene: popA
C:Superfamily: penicillin-binding protein 1B
C:Keywords: glycosyltransferase; hexosyltransferase
```

```
Query Match 99.5%; Score 3466; DB 2; Length 719;
Best Local Similarity 99.4%; Pred. No. 5.1e-193;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLKAYISIEDHFPDRGIDTRIIGAFLRN 60
|||||
DB 54 KIYDNKNQIADLGSERRVNAQANDIPTDLKAYISIEDHFPDRGIDTRIIGAFLRN 113
QY 61 LOSNSLOGGSLTQOLIKLTFYFSTSDQTSRKAQEWALAIQLEOKATKOEILTYINK 120
|||||
DB 114 LOSNSLOGGSLTQOLIKLTFYFSTSDQTSRKAQEWALAIQLEOKATKOEILTYINK 173
QY 121 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPQAPQYDPSHPEAQAORRNLYL 180
|||||
DB 174 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPQAPQYDPSHPEAQAORRNLYL 233
QY 181 SEMKNQGYISAEQYKAVNPITDGLQSLKASNPYAMNDVYLKEVINQVEEFGYNLLT 240
|||||
DB 234 SEMKNQGYISAEQYKAVNPITDGLQSLKASNPYAMNDVYLKEVINQVEEFGYNLLT 293
QY 241 TGMVYTNVDEQAKHLMIDYNTDEYVAYPDDELQVASTIVDVSNKGYIAQLGARHSSN 300
|||||
DB 294 TGMVYTNVDEQAKHLMIDYNTDEYVAYPDDELQVASTIVDVSNKGYIAQLGARHSSN 353
QY 301 VSEGINQAVETNRDQWGTMRPITDYAPALEGYVDSTATYHDEPYNPGTNTPYNNMR 360
|||||
DB 354 VSEGINQAVETNRDQWGTMRPITDYAPALEGYVDSTATYHDEPYNPGTNTPYNNMR 413
QY 361 GYFGNITLQYALQOOSRNPAVETLNKVGILNRAKTFNLGLIDYPSIHYSNAISSNTESD 420
```



```
|||||
Db 414 GYFGNTTLOALAOOSRNVAVETLNKVGIDRAKTFPLNGGIDIPSYHYSNAISSNTTESD 473
OY 421 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAMKETTAYAMTD 480
Db 474 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAMKETTAYAMTD 533
OY 481 MKKTVLTGTGRNAYLAMPPOAGKTGTSNVTDEIEINHKTISOFAVAPDELFACTYRKYSM 540
Db 534 MKKTVLTGTGRNAYLAMPPOAGKTGTSNVTDEIEINHKTISOFAVAPDELFACTYRKYSM 593
OY 541 AAWTGYSNRLTPLYGNGLTVAAKVYRSMTYIYSEGSNPEDMNPBGLYRNGEFVFNKGAR 600
Db 594 AAWTGYSNRLTPLYGNGLTVAAKVYRSMTYIYSEGSNPEDMNPBGLYRNGEFVFNKGAR 653
OY 601 STWNSPAPQOPPESTESSSSSDSTSSSTPTSTNNSTTTNNNTQOSONTPDOONON 660
Db 654 STWNSPAPQOPPESTESSSSSDSTSSSTPTSTNNSTTTNNNTQOSONTPDOONON 713
OY 661 POPAP 666
Db 714 POPAP 719
```

```
RESULT 6
S28034
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 56742) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28034
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of pen
A:Reference number: S28031; MUID:93010977
A:Accession: S28034
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67869; NID:947411; PIDN:CAA48069.1; PID:947412
C:Superfamily: penicillin-binding protein 1B
```

```
Query Match 93.2%; Score 3246; DB 2; Length 719;
Best Local Similarity 92.0%; Pred. No. 2.8e-180;
Matches 613; Conservative 31; Mismatches 22; Indels 0; Gaps 0;
OY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFDHGIDIRILGAFLRN 60
Db 54 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFDHGIDIRILGAFLRN 113
OY 61 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEWMLAIQLEQKATKQELTYIYINK 120
Db 114 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEWMLAIQLEQKATKQELTYIYINK 173
OY 121 VYMSNGNYGMOTRAQNYKYKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRNLVL 180
Db 174 VYMSNGNYGMOTRAQNYKYKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRNLVL 233
OY 181 SEKKNGYISAEQYKAVNTPITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLLT 240
Db 234 SEKKNGYISAEQYKAVNTPITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLLT 293
OY 241 TGMADVNTNDOEAKHLMIDYNTDEVAYAPDELAQVASTIVDVSNKVIQAOLGARHOSSN 300
Db 294 TGMADVNTNDOEAKHLMIDYNTDEVAYAPDELAQVASTIVDVSNKVIQAOLGARHOSSN 353
OY 301 VSEFGINQAVETNRDWSGSKKPTIDYAPALEYGYDSTATIVHDEPNYNGTNPVYNMNR 360
Db 354 VSEFGINQAVETNRDWSGSKKPTIDYAPALEYGYDSTATIVHDEPNYNGTNPVYNMNR 413
OY 361 GYFGNTTLOALAOOSRNVAVETLNKVGIDRAKTFPLNGGIDIPSYHYSNAISSNTTESD 420
Db 414 AYFGNTTLOALAOOSRNVAVETLNKVGIDRAKTFPLNGGIDIPSYHYSNAISSNTTESN 473
```

```
OY 421 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAMKETTAYAMTD 480
Db 474 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAMKETTAYAMTD 533
OY 481 MKKTVLTGTGRNAYLAMPPOAGKTGTSNVTDEIEINHKTISOFAVAPDELFACTYRKYSM 540
Db 534 MKKTVLTGTGRNAYLAMPPOAGKTGTSNVTDEIEINHKTISOFAVAPDELFACTYRKYSM 593
OY 541 AAWTGYSNRLTPLYGNGLTVAAKVYRSMTYIYSEGSNPEDMNPBGLYRNGEFVFNKGAR 600
Db 594 AAWTGYSNRLTPLYGNGLTVAAKVYRSMTYIYSEGSNPEDMNPBGLYRNGEFVFNKGAR 653
OY 601 STWNSPAPQOPPESTESSSSSDSTSSSTPTSTNNSTTTNNNTQOSONTPDOONON 660
Db 654 PWTERTSTQOOSTAESSSSDSTSSSTPTSTNNSTTTNNNTQOSONTPDOONON 713
OY 661 POPAP 666
Db 714 POPAP 719
```

```
RESULT 7
S28032
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 681) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28032
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of
A:Reference number: S28031; MUID:93010977
A:Accession: S28032
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67866; NID:947407; PIDN:CAA48066.1; PID:947408
C:Superfamily: penicillin-binding protein 1B
```

```
Query Match 93.1%; Score 3245; DB 2; Length 719;
Best Local Similarity 92.0%; Pred. No. 3.2e-180;
Matches 613; Conservative 35; Mismatches 18; Indels 0; Gaps 0;
OY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFDHGIDIRILGAFLRN 60
Db 54 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFDHGIDIRILGAFLRN 113
OY 61 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEWMLAIQLEQKATKQELTYIYINK 120
Db 114 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEWMLAIQLEQKATKQELTYIYINK 173
OY 121 VYMSNGNYGMOTRAQNYKYKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRNLVL 180
Db 174 VYMSNGNYGMOTRAQNYKYKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRNLVL 233
OY 181 SEKKNGYISAEQYKAVNTPITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLLT 240
Db 234 SEKKNGYISAEQYKAVNTPITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLLT 293
OY 241 TGMADVNTNDOEAKHLMIDYNTDEVAYAPDELAQVASTIVDVSNKVIQAOLGARHOSSN 300
Db 294 TGMADVNTNDOEAKHLMIDYNTDEVAYAPDELAQVASTIVDVSNKVIQAOLGARHOSSN 353
OY 301 VSEFGINQAVETNRDWSGSKKPTIDYAPALEYGYDSTATIVHDEPNYNGTNPVYNMNR 360
Db 354 VSEFGINQAVETNRDWSGSKKPTIDYAPALEYGYDSTATIVHDEPNYNGTNPVYNMNR 413
OY 361 GYFGNTTLOALAOOSRNVAVETLNKVGIDRAKTFPLNGGIDIPSYHYSNAISSNTTESD 420
Db 414 AYFGNTTLOALAOOSRNVAVETLNKVGIDRAKTFPLNGGIDIPSYHYSNAISSNTTESN 473
OY 421 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAMKETTAYAMTD 480
```



```
Db 474 KÖYGASSEKMAAAYAAAFANGGIYHNPYINKVYFSDGSKKEFSVDGTFRAMKETTAAYMTE 533
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 481 MKKTYLTGTRNAYLAWLPQAGKTGTSNYTDEIEENHIKTSQFVAPDELPAGYTRKYSM 540
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 534 MKKTYLTGTRGAGVLPPLPQAGKTGTSNYTDEIEENHIKNTGYAPDEPMVGTTRKYSM 593
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 541 AVMTGYSNRLTPLVGNGITVAAKYRSMNTYLSGSGNPDENNIEGLYRNGEYFVKNGAR 600
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 594 AVMTGYSNRLTPIVGDDGLVAAKYRSMITLYSEDTHEPDMTPDGLFRNGEYFVKNGAR 653
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 601 STWNSPAPQPPSTESSSSSDSSSTSSSTTPSTNNSTTNNPNNNTQOQSNTPDQOQON 660
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 624 STWNSPAPQPPSTESSSSSDSSSTSSSTTPSTNNSTTNNPNNNTQOQSNTPDQOQON 713
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 661 POPAP 666
|||||
Db 714 POPAP 719
```

RESULT 8

S28031

penicillin-binding protein 1a - Streptococcus pneumoniae (strain 456) (fragment)

C:Species: Streptococcus pneumoniae

C>Date: 17-Apr-1993 #sequence.revision 17-Apr-1993 #text-change 26-May-2000

C/Accession: S28031

R:Martin, C.; Sibbold, C.; Hakenbeck, R.

EMBO J. 11, 3831-3836, 1992

A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of per

A:Reference number: S28031; MUID:93010977

A:Accession: S28031

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-719 <MAR>

A:Cross-references: EMBL:X67866; NID:947405; PIDN:CAA48066.1; PID:947406

C:Superfamily: penicillin-binding protein 1B

```
Query Match 93.1%; Score 3243; DB 2; Length 719;
Best Local Similarity 91.9%; Pred. No. 4.2e-180;
Matches 612; Conservative 32; Mismatches 22; Indels 0; Gaps 0;

Oy 1 KIYDNKNQDLADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 54 KIYDNKNQDLADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 61 LOSNSLOGSSTLTQOLKLTFTSTSDQTSRKAQEAALAIQLEOKATKQEIILTYINK 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 114 LOSNSLOGSSTLTQOLKLTFTSTSDQTSRKAQEAALAIQLEOKATKQEIILTYINK 173
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 121 VYMSNGNMGOTAAQNYGKDLNLSLPQALLAGMPAPQOYDPSHPEAAQDRNLVL 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 174 VYMSNGNMGOTAAQNYGKDLNLSLPQALLAGMPAPQOYDPSHPEAAQDRNLVL 233
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 181 SEMKNQGYISAEOYEKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYNLT 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 234 SEMKNQGYISAEOYEKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYNLT 293
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 241 TGMVYTNVDEAOKHLMIDINTDEYVAYPDDELQVASTIVDVNSGKYIAOLGARHOSN 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 294 TGMVYTNVDEAOKHLMIDINTDEYVAYPDDELQVASTIVDVNSGKYIAOLGARHOSN 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 301 VSGFNQAVETNRDMGSMKPIITDYAPALEGYVDSTATTIYHDEBYNPGNTPYNMNR 360
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 354 VSGFNQAVETNRDMGSMKPIITDYAPALEGYVDSTATTIYHDEBYNPGNTPYNMNR 413
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 361 GYFNGITLQVLAQOOSRNYPAVETLNKVLNRAKTFNLGIDIDPSIHNSMAISSNTTESD 420
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 414 AYFGNITLQVLAQOOSRNYPAVETLNKVLNRAKTFNLGIDIDPSIHNSMAISSNTTESN 473
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 421 KKYGASSEKMAAAYAAAFANGSTYYKPMYIHKVYFSDGSEKFEFSNVGTFRAMKETTAAYMTD 480
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 474 KÖYGASSEKMAAAYAAAFANGSTYYKPMYIHKVYFSDGSEKFEFSNVGTFRAMKETTAAYMTD 533
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Oy 481 MKKTYLTGTRNAYLAWLPQAGKTGTSNYTDEIEENHIKTSQFVAPDELPAGYTRKYSM 540
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 534 MKKTYLTGTRGAGVLPPLPQAGKTGTSNYTDEIEENHIKNTGYAPDEPMVGTTRKYSM 593
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 541 AVMTGYSNRLTPLVGNGITVAAKYRSMNTYLSGSGNPDENNIEGLYRNGEYFVKNGAR 600
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 594 AVMTGYSNRLTPIVGDDGLVAAKYRSMITLYSEDTHEPDMTPDGLFRNGEYFVKNGAR 653
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 601 STWNSPAPQPPSTESSSSSDSSSTSSSTTPSTNNSTTNNPNNNTQOQSNTPDQOQON 660
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 624 STWNSPAPQPPSTESSSSSDSSSTSSSTTPSTNNSTTNNPNNNTQOQSNTPDQOQON 713
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 661 POPAP 666
|||||
Db 714 POPAP 719
```

RESULT 9

S28033

penicillin-binding protein 1a - Streptococcus pneumoniae (strain 670) (fragment)

C:Species: Streptococcus pneumoniae

C>Date: 17-Apr-1993 #sequence.revision 17-Apr-1993 #text-change 26-May-2000

C/Accession: S28033

R:Martin, C.; Sibbold, C.; Hakenbeck, R.

EMBO J. 11, 3831-3836, 1992

A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of

A:Reference number: S28031; MUID:93010977

A:Accession: S28033

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-719 <MAR>

A:Cross-references: EMBL:X67867; NID:947409; PIDN:CAA48067.1; PID:947410

C:Superfamily: penicillin-binding protein 1B

```
Query Match 93.0%; Score 3241; DB 2; Length 719;
Best Local Similarity 91.9%; Pred. No. 5.5e-180;
Matches 612; Conservative 36; Mismatches 18; Indels 0; Gaps 0;

Oy 1 KIYDNKNQDLADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 54 KIYDNKNQDLADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 61 LOSNSLOGSSTLTQOLKLTFTSTSDQTSRKAQEAALAIQLEOKATKQEIILTYINK 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 114 LOSNSLOGSSTLTQOLKLTFTSTSDQTSRKAQEAALAIQLEOKATKQEIILTYINK 173
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 121 VYMSNGNMGOTAAQNYGKDLNLSLPQALLAGMPAPQOYDPSHPEAAQDRNLVL 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 174 VYMSNGNMGOTAAQNYGKDLNLSLPQALLAGMPAPQOYDPSHPEAAQDRNLVL 233
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 181 SEMKNQGYISAEOYEKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYNLT 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 234 SEMKNQGYISAEOYEKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYNLT 293
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 241 TGMVYTNVDEAOKHLMIDINTDEYVAYPDDELQVASTIVDVNSGKYIAOLGARHOSN 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 294 TGMVYTNVDEAOKHLMIDINTDEYVAYPDDELQVASTIVDVNSGKYIAOLGARHOSN 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 301 VSGFNQAVETNRDMGSMKPIITDYAPALEGYVDSTATTIYHDEBYNPGNTPYNMNR 360
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 354 VSGFNQAVETNRDMGSMKPIITDYAPALEGYVDSTATTIYHDEBYNPGNTPYNMNR 413
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 361 GYFNGITLQVLAQOOSRNYPAVETLNKVLNRAKTFNLGIDIDPSIHNSMAISSNTTESD 420
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 414 AYFGNITLQVLAQOOSRNYPAVETLNKVLNRAKTFNLGIDIDPSIHNSMAISSNTTESN 473
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 421 KKYGASSEKMAAAYAAAFANGSTYYKPMYIHKVYFSDGSEKFEFSNVGTFRAMKETTAAYMTD 480
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 474 KÖYGASSEKMAAAYAAAFANGSTYYKPMYIHKVYFSDGSEKFEFSNVGTFRAMKETTAAYMTE 533
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 481 MKKTYLTGTRNAYLAWLPQAGKTGTSNYTDEIEENHIKTSQFVAPDELPAGYTRKYSM 540
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```


Db 534 MKKTVLVYIGRGAYLFWLPQAGKTGTSNYDEIEKYIKMTGYVAPDEMFVGYTRKYAM 593
QY 541 AVMGTYSNRLTPWVG 555
|||||
Db 594 AVMGTYSNRLTPWVG 608

RESULT 12

B42893 penicillin-binding protein 1 - Streptococcus oralis

C:Species: Streptococcus oralis

C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 04-Mar-2000

C:Accession: B42893

R:Martin, C.; Briesse, T.; Hakenbeck, R.

J. Bacteriol. 174, 4517-4523, 1992

A:Title: Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus oralis

A:Reference number: A42893; MUID:92325042

A:Accession: B42893

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-637 <MAR>

A:Cross-references: GB:M90528; NID:g153769; PIDN:AAA26958.1; PID:g153771

C:Superfamily: penicillin-binding protein 1B

Query Match 77.5%; Score 2698.5; DB 2; Length 637;

Best Local Similarity 87.5%; Pred. No. 1.2e-148;

Matches 511; Conservative 39; Mismatches 33; Indels 1; Gaps 1;

QY 1 KIYDNKQIADLGSERRVNAQANDIPDLVKAIVSIEDHREDFRGIDTIRILGAFRLN 60
Db 54 KIYDNKQIADLGSERRVNAQANDIPDLVKAIVSIEDHREDFRGIDTIRILGAFRLN 113
QY 61 LQ-SNSLOGSSTLQOLIKLTYFSTSDQTSRKQAEMLAQLQKATKQELITTYIN 119
Db 114 LRGGGLQAGSTLQOLIKLTYFSTSDQTSRKQAEMLAQLQKATKQELITTYIN 173
QY 120 KYVMSNGVGMOTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPPEAODRRNLV 179
Db 174 KYVMSNGVGMOTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPPEAODRRNLV 233
QY 180 LSEMKNOGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNYLKEVINOVEEETGYNLT 239
Db 234 LSEMKNOGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNYLKEVINOVEEETGYNLT 293
QY 240 TTGMADYTNVDOEAKHMDIYNTDEYVAYPPDELOVASTYDVSNKGYIAQLGARHOS 299
Db 294 TTGMADYTNVDOEAKHMDIYNTDEYVAYPPDELOVASTYDVSNKGYIAQLGARHOS 353
QY 360 NVSGINQAVETNRDMSGTMRPTIDYAPALEYGYDSTATYVDEHPYVPGNTPTVYVMD 359
Db 354 NVSGINQAVETNRDMSGTMRPTIDYAPALEYGYDSTATYVDEHPYVPGNTPTVYVMD 413
QY 360 RGYEGNITLQYALQOSRNVPAVETLNKVGILNRAKTFNLGLIDYPSIHYSAISSNTTES 419
Db 414 KSYEGNITLQYALQOSRNVPAVETLNKVGILNRAKTFNLGLIDYPSIHYSAISSNTTES 473
QY 420 DKRTGASSEKMAAAYAAAFANGSTYKPMYIHKYVPSDGESEKESFNGVTRAKETAYMMT 479
Db 474 DKRTGASSEKMAAAYAAAFANGSTYKPMYIHKYVPSDGESEKESFNGVTRAKETAYMMT 533
QY 480 DMKTVLVLTGTRNAYLAWLPOAGKTGTSNYTDEIEHNTSOPVAPDELFAGYTRKYS 539
Db 534 DMKTVLVLTGTRNAYLAWLPOAGKTGTSNYTDEIEHNTSOPVAPDELFAGYTRKYS 593
QY 540 MAVMTGYSNRLTPWVGNGLTVAAYKYSRMTYLSGNSPDEDMNI 583
Db 594 MAVMTGYSNRLTPWVGNGLTVAAYKYSRMTYLSGNSPDEDMNI 637

RESULT 13

G86692

penicillin-binding protein 1A [imported] - Lactococcus lactis subsp. lactis (strain I
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: G86692
R:Boletín, A.; Winkler, P.; Mager, S.; Jallón, O.; Malame, K.; Weissbach, J.; Eh
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: G86692
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-664 <STO>
A:Cross-references: GB:AE005176; PID:g12723428; PIDN:AAK04641.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ponA
C:Superfamily: penicillin-binding protein 1B

Query Match 47.6%; Score 1657.5; DB 2; Length 664;
Best Local Similarity 52.1%; Pred. No. 2.6e-88;

Matches 333; Conservative 105; Mismatches 166; Indels 35; Gaps 6;

QY 1 KIYDNKQIADLGSERRVNAQANDIPDLVKAIVSIEDHREDFRGIDTIRILGAFRLN 60
Db 49 KYVDSERKVAATLGAEOENLVKTDNIPVWLVAVNSIEDHREDFRGIDTIRILGAFVNN 108
QY 61 LQSNLSLOGSSTLQOLIKLTYFSTSDQTSRKQAEMLAQLQKATKQELITTYIN 120
Db 109 LRGGGLQAGSTLQOLIKLTYFSTSDQTSRKQAEMLAQLQKATKQELITTYIN 168
QY 121 KYVMSNGVGMOTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPPEAODRRNLV 180
Db 169 KYVMSNGVGMOTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPPEAODRRNLV 228
QY 241 TTGMADYTNVDOEAKHMDIYNTDEYVAYPPDELOVASTYDVSNKGYIAQLGARHOS 300
Db 289 TTGMADYTNVDOEAKHMDIYNTDEYVAYPPDELOVASTYDVSNKGYIAQLGARHOS 348
QY 301 VSGINQAVETNRDMSGTMRPTIDYAPALEYGYDSTATYVDEHPYVPGNTPTVYVMD 360
Db 349 VSGINQAVETNRDMSGTMRPTIDYAPALEYGYDSTATYVDEHPYVPGNTPTVYVMD 407
QY 361 RGYEGNITLQYALQOSRNVPAVETLNKVGILNRAKTFNLGLIDYPSIHYSAISSNTTES 420
Db 408 RGYEGNITLQYALQOSRNVPAVETLNKVGILNRAKTFNLGLIDYPSIHYSAISSNTTES 463
QY 421 KRTGASSEKMAAAYAAAFANGSTYKPMYIHKYVPSDGESEKESFNGVTRAKETAYMMT 480
Db 464 KRTGASSEKMAAAYAAAFANGSTYKPMYIHKYVPSDGESEKESFNGVTRAKETAYMMT 523
QY 481 DMKTVLVLTGTRNAYLAWLPOAGKTGTSNYTDEIEHNTSOPVAPDELFAGYTRKYS 532
Db 524 DMKTVLVLTGTRNAYLAWLPOAGKTGTSNYTDEIEHNTSOPVAPDELFAGYTRKYS 583
QY 533 GYTRKYSNAVMTGYSNRLTPWVGNGLTVAAYKYSRMTYLSGNSPDEDMNI 591
Db 584 GYTRKYSNAVMTGYSNRLTPWVGNGLTVAAYKYSRMTYLSGNSPDEDMNI 643
QY 592 EFVFNKARSTWNSPAPQOPSTESSSSSDSSTQSSS 630
Db 644 TALVK-----TDSGQTTSSQSSA 661

RESULT 14

AD1311

penicillin-binding protein 2A homolog pbpa [imported] - Listeria monocytogenes (strai
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

Fri Jun 14 10:20:22 2002

us-08-961-083-2.rpr

Page 9

Db 709 AEDKKTEEEKKEEAK 727

Search completed: June 13, 2002, 08:41:08
Job time: 34 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:41:09 ; Search time 34.23 Seconds
(without alignments)
3365.896 Million cell updates/sec

Title: US-08-961-083-2
Perfect score: 3484
Sequence: 1 KIYDNKNQIADLGSERRVN.....TQSNMTPDQGNQNPQAP 666

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._viral:*
16: sp._bacteriophage:*
17: sp._archaeoprotein:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3484	100.0	719	2	Q9REU0 streptococc
2	3468	99.5	719	2	Q9REU9 streptococc
3	3403	97.7	719	2	Q9REU3 streptococc
4	3329	95.6	719	2	Q9WVW0 streptococc
5	3314	95.1	719	2	Q9REU8 streptococc
6	3299	94.7	719	2	Q9WVW1 streptococc
7	3274	94.0	719	2	Q9REU4 streptococc
8	3247	93.2	719	2	Q9REU7 streptococc
9	3246	93.2	719	2	Q9REU6 streptococc
10	3245	93.2	719	2	Q9REU5 streptococc
11	3245	93.1	719	2	Q9REU4 streptococc
12	3243	93.1	719	2	Q9REU5 streptococc
13	3241	93.0	719	2	Q9REU6 streptococc
14	3231	92.7	719	2	Q9REU5 streptococc
15	3192	91.6	719	2	Q9REU5 streptococc
16	3134	90.0	718	2	Q9REU7 streptococc

17	2820	80.9	608	2	Q54951 streptococc
18	2791	80.1	728	2	Q9F267 streptococc
19	2696	77.4	727	2	Q9F264 streptococc
20	2432	69.8	527	2	Q70037 streptococc
21	2091	60.0	398	2	Q9R7M6 streptococc
22	2088	59.9	398	2	Q87102 streptococc
23	2037.5	58.5	721	16	Q99YX1 streptococc
24	2029	58.2	398	2	Q9R7M5 streptococc
25	2015	57.8	398	2	Q87108 streptococc
26	2011	57.7	398	2	Q87105 streptococc
27	1936	55.6	398	2	Q9R7M7 streptococc
28	1926	55.3	398	2	Q87103 streptococc
29	1921	55.1	398	2	Q9R7M8 streptococc
30	1911	54.9	793	2	Q9EXM9 streptococc
31	1881	54.0	398	2	Q87107 streptococc
32	1876	53.8	398	2	Q87104 streptococc
33	1757	50.4	400	2	Q87106 streptococc
34	1751	50.3	778	2	Q9EXN1 streptococc
35	1657.5	47.6	664	16	Q9C123 streptococc
36	1592	45.7	310	2	Q52741 streptococc
37	1583	45.4	310	2	Q52743 streptococc
38	1573	45.1	310	2	Q52739 streptococc
39	1571	45.1	310	2	Q52737 streptococc
40	1567	45.0	310	2	Q52738 streptococc
41	1564	44.9	310	2	Q52740 streptococc
42	1558	44.7	310	2	Q52742 streptococc
43	1454	41.7	310	2	Q93N72 streptococc
44	1449	41.6	310	2	Q52744 streptococc
45	1201	34.5	826	16	Q92AB6 streptococc

ALIGNMENTS

RESULT 1
Q9REU0 PRELIMINARY; PRT; 719 AA.
AC Q9REU0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=SP 1261;
RX MEDLINE=21432820; PubMed=11549185;
RA Ferroni A., Berche P.;
RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
RT penicillin-resistant clinical isolates of Streptococcus pneumoniae
RT serotype 23F from the nasopharyngeal flora of children.";
RT J. Med. Microbiol. 50:828-832(2001).
RL EMBL: AF210745; AAF17255.1; -;
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
SO SEQUENCE 719 AA; 79758 MW; 5BD397EB3B4B3AA6 CRC64;

Query Match 100.0%; Score 3484; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 2, 1e-181;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFPDHRGIDTIRIIGAFIRN 60
DB 54 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFPDHRGIDTIRIIGAFIRN 113


```

QY 61 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKRKAQEWLAIOLEKATKQELITYYINK 120
    |||||||
Db 114 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKRKAQEWLAIOLEKATKQELITYYINK 173
QY 121 VYMSNGNYGMQTRAQNYTGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRNLV 180
    |||||||
Db 174 VYMSNGNYGMQTRAQNYTGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRNLV 233
QY 181 SEMKNGYISAEQYERKAVNTPITDGLQSLKSASNPAYMDNYLKEYINQVEETGYNLLT 240
    |||||||
Db 234 SEMKNGYISAEQYERKAVNTPITDGLQSLKSASNPAYMDNYLKEYINQVEETGYNLLT 293
QY 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVYAYPDELOVASTIVDYSNGKVIAOLGARHQQSN 300
    |||||||
Db 294 TGMDEVYTNVDOEAQKHLMDIYNTDEVYAYPDELOVASTIVDYSNGKVIAOLGARHQQSN 353
QY 301 VSRFGINQAVETNRDMSGTKPITDYAPALEGYGYDSTATVHDEPNYNGTNPVYNMNR 360
    |||||||
Db 354 VSRFGINQAVETNRDMSGTKPITDYAPALEGYGYDSTATVHDEPNYNGTNPVYNMNR 413
QY 361 GYFGNITLQYALQOQRNVPAVETLNKVGLNRAKTFNLGIDIPSYHNSAISSNTTESD 420
    |||||||
Db 414 GYFGNITLQYALQOQRNVPAVETLNKVGLNRAKTFNLGIDIPSYHNSAISSNTTESD 473
QY 421 KRYGASSEKMAAAYAFANGGTYKPMYTHKVFSDGSEKEFSNVGTRAKETAYAMTD 480
    |||||||
Db 474 KRYGASSEKMAAAYAFANGGTYKPMYTHKVFSDGSEKEFSNVGTRAKETAYAMTD 533
QY 481 MKTIVLTGTGRNAYLAMLPOAGKTGTSNTTDEIENHIKTSQFVAPDELFACTYKRYSM 540
    |||||||
Db 534 MKTIVLTGTGRNAYLAMLPOAGKTGTSNTTDEIENHIKTSQFVAPDELFACTYKRYSM 593
QY 541 AVWTGYSNRLTPLYNGCLTYAAKVYRSMTYLSGGSNPEDMNIPEGLYRNGEYVFNKGAR 600
    |||||||
Db 594 AVWTGYSNRLTPLYNGCLTYAAKVYRSMTYLSGGSNPEDMNIPEGLYRNGEYVFNKGAR 653
QY 601 STWNSPAPQOPSTESSSSSDSSTSSQSSSTPTSTNNSTTTNNNTQOQSNTPPDOQON 660
    |||||||
Db 654 STWNSPAPQOPSTESSSSSDSSTSSQSSSTPTSTNNSTTTNNNTQOQSNTPPDOQON 713
QY 661 POPAP 666
    |||||||
Db 714 POPAP 719

```

```

SQ SEQUENCE 719 AA: 79830 MW: 5f67768DEAE3840 CRC64:
Query Match 99.5%; Score 3468; DB 2; Length 719;
Best Local Similarity 99.5%; Pred. No. 1.6e-180;
Matches 663; Conservative 2; Mismatches 1; Indels 0; Gaps 0:
QY 1 KIDNKNOLIADIGSERRRVNAQANDIPTDLVKAIVSIEDHRFPDHGIDITRILGAFLRN 60
    |||||||
Db 54 KIDNKNOLIADIGSERRRVNAQANDIPTDLVKAIVSIEDHRFPDHGIDITRILGAFLRN 113
QY 61 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKRKAQEWLAIOLEKATKQELITYYINK 120
    |||||||
Db 114 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKRKAQEWLAIOLEKATKQELITYYINK 173
QY 121 VYMSNGNYGMQTRAQNYTGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRNLV 180
    |||||||
Db 174 VYMSNGNYGMQTRAQNYTGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRNLV 233
QY 181 SEMKNGYISAEQYERKAVNTPITDGLQSLKSASNPAYMDNYLKEYINQVEETGYNLLT 240
    |||||||
Db 234 SEMKNGYISAEQYERKAVNTPITDGLQSLKSASNPAYMDNYLKEYINQVEETGYNLLT 293
QY 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVYAYPDELOVASTIVDYSNGKVIAOLGARHQQSN 300
    |||||||
Db 294 TGMDEVYTNVDOEAQKHLMDIYNTDEVYAYPDELOVASTIVDYSNGKVIAOLGARHQQSN 353
QY 301 VSRFGINQAVETNRDMSGTKPITDYAPALEGYGYDSTATVHDEPNYNGTNPVYNMNR 360
    |||||||
Db 354 VSRFGINQAVETNRDMSGTKPITDYAPALEGYGYDSTATVHDEPNYNGTNPVYNMNR 413
QY 361 GYFGNITLQYALQOQRNVPAVETLNKVGLNRAKTFNLGIDIPSYHNSAISSNTTESD 420
    |||||||
Db 414 GYFGNITLQYALQOQRNVPAVETLNKVGLNRAKTFNLGIDIPSYHNSAISSNTTESD 473
QY 421 KRYGASSEKMAAAYAFANGGTYKPMYTHKVFSDGSEKEFSNVGTRAKETAYAMTD 480
    |||||||
Db 474 KRYGASSEKMAAAYAFANGGTYKPMYTHKVFSDGSEKEFSNVGTRAKETAYAMTD 533
QY 481 MKTIVLTGTGRNAYLAMLPOAGKTGTSNTTDEIENHIKTSQFVAPDELFACTYKRYSM 540
    |||||||
Db 534 MKTIVLTGTGRNAYLAMLPOAGKTGTSNTTDEIENHIKTSQFVAPDELFACTYKRYSM 593
QY 541 AVWTGYSNRLTPLYNGCLTYAAKVYRSMTYLSGGSNPEDMNIPEGLYRNGEYVFNKGAR 600
    |||||||
Db 594 AVWTGYSNRLTPLYNGCLTYAAKVYRSMTYLSGGSNPEDMNIPEGLYRNGEYVFNKGAR 653
QY 601 STWNSPAPQOPSTESSSSSDSSTSSQSSSTPTSTNNSTTTNNNTQOQSNTPPDOQON 660
    |||||||
Db 654 STWNSPAPQOPSTESSSSSDSSTSSQSSSTPTSTNNSTTTNNNTQOQSNTPPDOQON 713
QY 661 POPAP 666
    |||||||
Db 714 POPAP 719

```

```

RESULT 2
O9RET9 PRELIMINARY: PRT: 719 AA.
AC O9RET9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP 1513;
RX MEDLINE=21432820; PubMed=11549185;
RA Ferroni A., Berche P.;
RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
penicillin-resistant clinical isolates of Streptococcus pneumoniae
serotype 23F from the nasopharyngeal flora of children.";
RL J. Med. Microbiol. 50:828-832(2001).
DR EMBL; AF210746; AAF17256.1;
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00903; Transpeptidase; 1.
DR PDBom: PDB01895; Transglycosyl; 1.

```

```

RESULT 3
O9RET3 PRELIMINARY: PRT: 719 AA.
AC O9RET3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BM 4200;
RX MEDLINE=21432820; PubMed=11549185;

```


RA Ferroni A., Berche P.;
RT "Alterations to penicillin-binding proteins 1a, 2b and 2x amongst
penicillin-resistant clinical isolates of Streptococcus pneumoniae
serotype 23F from the nasopharyngeal flora of children.";
RL J. Med. Microbiol. 50:828-832(2001).
DR EMBL: AF210752; AAF17262.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl. 1.
DR Pfam: PF00905; Transpeptidase. 1.
DR Prodom: PD001895; Transglycosyl. 1.
SQ SEQUENCE 719 AA; 79781 MW; DB5993F6BD47F72D CRC64;

Query Match 97.7%; Score 3403; DB 2; Length 719;
Best Local Similarity 97.3%; Pred. No. 5.4e-177;
Matches 648; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 KIYDNKNLIDLGSERRVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFRLN 60
DB 54 KIYDNKNLIDLGSERRVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFRLN 113
QY 61 LOSNSLOGSSTLTOOLIKLTYFSTSTSDQTSRKAEAMLAIOLEOKATKOEILTYINK 120
DB 114 LOSNSLOGSSTLTOOLIKLTYFSTSTSDQTSRKAEAMLAIOLEOKATKOEILTYINK 173
QY 121 VYMSNGNYGMOPTAONNYGKDLNNLSLPOLALLGMPAPAPNOYDPSHPEAODRRNLVL 180
DB 174 VYMSNGNYGMOPTAONNYGKDLNNLSLPOLALLGMPAPAPNOYDPSHPEAODRRNLVL 233
QY 181 SEMKNOGYISAEOYEKAVNTPITDGLSLKSASNPAYMDNYLKEVINOVEETGYNLLT 240
DB 234 SEMKNOGYISAEOYEKAVNTPITDGLSLKSASNPAYMDNYLKEVINOVEETGYNLLT 293
QY 241 TGMVYTNVDOEAKHMLDINTDEYVAYPDDELQVASTIVDVNSGKVIQALGARHOSN 300
DB 294 TGMVYTNVDOEAKHMLDINTDEYVAYPDDELQVASTIVDVNSGKVIQALGARHOSN 353
QY 301 VSFGINQAVENNRDMGSMKPTITDYAPALEGYVDSTATTIVHDEPYNPGTNPYYNMDR 360
DB 354 VSFGINQAVENNRDMGSMKPTITDYAPALEGYVDSTATTIVHDEPYNPGTNPYYNMDR 413
QY 361 GYFNGITLOVALOOSRNPVAVETLNKVLNRAKTFPLNGLDIDYPSIHNSNAISSWTESD 420
DB 414 GYFNGITLOVALOOSRNPVAVETLNKVLNRAKTFPLNGLDIDYPSIHNSNAISSWTESD 473
QY 421 KRYGASSEKMAAAYAAAFANGSTYKPMYIHKVVFSDGSEKEFSNVGTRAMKETTAYMMD 480
DB 474 KRYGASSEKMAAAYAAAFANGSTYKPMYIHKVVFSDGSEKEFSNVGTRAMKETTAYMMD 533
QY 481 MKKTVLYTGGRNAYLAMLPOAGKTGTSNYTDEIEHNIKTSQFVAPDELFACTRKYSM 540
DB 534 MKKTVLYTGGRNAYLAMLPOAGKTGTSNYTDEIEHNIKTSQFVAPDELFACTRKYSM 593
QY 541 AVMTGYSNRLTPPLVGNGLTVAAKYRSMMTYLSGSPNEDNINPEGLRNGEFPVKNGAR 600
DB 594 AVMTGYSNRLTPPLVGNGLTVAAKYRSMMTYLSGSPNEDNINPEGLRNGEFPVKNGAR 653
QY 601 STWNSPAPQPPSTRESSSSSDSSTSPSTNNSTTNPNNNTQOQSTTTPDOQON 660
DB 654 STWNSPAPQPPSTRESSSSSDSSTSPSTNNSTTNPNNNTQOQSTTTPDOQON 713
QY 661 POPAP 666
DB 714 POPAP 719

RESULT 4
Q9MYW0 PRELIMINARY; PRT: 719 AA.
AC Q9MYW0;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)

DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
ON NCBI_TaxID=1313;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PO-342, PO-273, AND PO-341;
KA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
clone of Streptococcus pneumoniae arose by large recombinational
RT replacements of the cpsA-pbpA region.";
RL Microbiology 0:0-0(1999)
DR EMBL: AF139886; AAD43069.1; -
DR EMBL: AF139884; AAD43067.1; -
DR EMBL: AF139885; AAD43068.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl. 1.
DR Pfam: PF00905; Transpeptidase. 1.
DR Prodom: PD001895; Transglycosyl. 1.
SQ SEQUENCE 719 AA; 79622 MW; 7EBA6A75EA8FF8B3 CRC64;

Query Match 95.6%; Score 3329; DB 2; Length 719;
Best Local Similarity 94.4%; Pred. No. 5.6e-173;
Matches 629; Conservative 25; Mismatches 12; Indels 0; Gaps 0;

QY 1 KIYDNKNLIDLGSERRVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFRLN 60
DB 54 KIYDNKNLIDLGSERRVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFRLN 113
QY 61 LOSNSLOGSSTLTOOLIKLTYFSTSTSDQTSRKAEAMLAIOLEOKATKOEILTYINK 120
DB 114 LOSNSLOGSSTLTOOLIKLTYFSTSTSDQTSRKAEAMLAIOLEOKATKOEILTYINK 173
QY 121 VYMSNGNYGMOPTAONNYGKDLNNLSLPOLALLGMPAPAPNOYDPSHPEAODRRNLVL 180
DB 174 VYMSNGNYGMOPTAONNYGKDLNNLSLPOLALLGMPAPAPNOYDPSHPEAODRRNLVL 233
QY 181 SEMKNOGYISAEOYEKAVNTPITDGLSLKSASNPAYMDNYLKEVINOVEETGYNLLT 240
DB 234 SEMKNOGYISAEOYEKAVNTPITDGLSLKSASNPAYMDNYLKEVINOVEETGYNLLT 293
QY 241 TGMVYTNVDOEAKHMLDINTDEYVAYPDDELQVASTIVDVNSGKVIQALGARHOSN 300
DB 294 TGMVYTNVDOEAKHMLDINTDEYVAYPDDELQVASTIVDVNSGKVIQALGARHOSN 353
QY 301 VSFGINQAVENNRDMGSMKPTITDYAPALEGYVDSTATTIVHDEPYNPGTNPYYNMDR 360
DB 354 VSFGINQAVENNRDMGSMKPTITDYAPALEGYVDSTATTIVHDEPYNPGTNPYYNMDR 413
QY 361 GYFNGITLOVALOOSRNPVAVETLNKVLNRAKTFPLNGLDIDYPSIHNSNAISSWTESD 420
DB 414 GYFNGITLOVALOOSRNPVAVETLNKVLNRAKTFPLNGLDIDYPSIHNSNAISSWTESD 473
QY 421 KRYGASSEKMAAAYAAAFANGSTYKPMYIHKVVFSDGSEKEFSNVGTRAMKETTAYMMD 480
DB 474 KRYGASSEKMAAAYAAAFANGSTYKPMYIHKVVFSDGSEKEFSNVGTRAMKETTAYMMD 533
QY 481 MKKTVLYTGGRNAYLAMLPOAGKTGTSNYTDEIEHNIKTSQFVAPDELFACTRKYSM 540
DB 534 MKKTVLYTGGRNAYLAMLPOAGKTGTSNYTDEIEHNIKTSQFVAPDELFACTRKYSM 593
QY 541 AVMTGYSNRLTPPLVGNGLTVAAKYRSMMTYLSGSPNEDNINPEGLRNGEFPVKNGAR 600
DB 594 AVMTGYSNRLTPPLVGNGLTVAAKYRSMMTYLSGSPNEDNINPEGLRNGEFPVKNGAR 653
QY 601 STWNSPAPQPPSTRESSSSSDSSTSPSTNNSTTNPNNNTQOQSTTTPDOQON 660
DB 654 STWNSPAPQPPSTRESSSSSDSSTSPSTNNSTTNPNNNTQOQSTTTPDOQON 713

OY 661 POPAP 666
 Db 714 POPAP 719

RESULT 5

O9RET8 PRELIMINARY: PRT: 719 AA.

AC O9RET8: 01-MAY-2000 (TREMREL. 13, Created)
 DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
 DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBPIA.

OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OX NCBI_TaxID=1313;

RP SEQUENCE FROM N.A.

RC STRAIN-SP 1465;

RA MEDLINE=21432820; PubMed=11549185;

RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst

penicillin-resistant clinical isolates of Streptococcus pneumoniae

serotype 23F from the nasopharyngeal flora of children."

RU J. Med. Microbiol. 50:828-832(2001).

DR EMBL, AF210747; AAF17257.1; -

DR InterPro: IPR001264; Transglycosyl.

DR InterPro: IPR001460; Transpeptidase.

DR Pfam: PF00912; Transglycosyl; 1.

DR Pfam: PF00905; Transpeptidase; 1.

DR Prodom: PD001895; Transglycosyl; 1.

SQ SEQUENCE 719 AA; 79662 MW; A7B96EA15434A805 CRC64;

Query Match 95.1%; Score 3314; DB 2; Length 719;
 Best Local Similarity 94.1%; Pred. No. 3.7e-172;
 Matches 627; Conservative 27; Mismatches 12; Indels 0; Gaps 0;

OY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
 Db 54 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
 OY 61 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSRKAQEWLAIOLEKATKOEILTYINK 120
 Db 114 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSRKAQEWLAIOLEKATKOEILTYINK 173
 OY 121 VYMSNGNYGMQTAQNYKYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPBAADRNLVY 180
 Db 174 VYMSNGNYGMQTAQNYKYGKDLRELISLPOLALLAGMPQAPNOYDPYSHPBAADRNLVY 233
 OY 181 SEMKNGYISAQYKAVNTPTDGLQSLKASNNYPAVMDNYLKEYINOVEESTGNLIT 240
 Db 234 SEMKGGYISAQYKAVNTPTDGLQSLKSVNSYPAYMDNYLKEYIDVEGETGNLIT 293
 OY 241 TGMADYTYNDQEAOKHLMIDYNTDEVVAYAPDELOVASTIVDSNGKVAIOLGARQSSN 300
 Db 294 TGMADYTYNDQEAOKHLMIDYNTDEVVAYAPDELOVASTIVDSNGKVAIOLGARQSSN 353
 OY 301 VSEGINQAVETNRDMGSTMKPTTDVAPALEYGVYDSTATIVHDEPNYEGTNTPVYNNMR 360
 Db 354 VSEGINQAVETNRDMGSTMKPTTDVAPALEYGVYDSTATIVHDEPNYEGTNTPVYNNMR 413
 OY 361 GYFGNITLQYALQOSRNVPVETLKNVGLNRAKTEFLNGGIDYPSHYNSAISTNTESD 420
 Db 414 AVEFGNITLQYALQOSRNVPVETLKNVGLNRAKTEFLNGGIDYPSHYNSAISTNTESD 473
 OY 421 KYVGASSEKMAAAYAFANAGTYTKPMYTHKVEPSDSEKESPNVETRAMKETATAMMD 480
 Db 474 KYVGASSEKMAAAYAFANAGTYTKPMYTHKVEPSDSEKESPNVETRAMKETATAMMD 533

OY 481 MKKTVLTGCTGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 540
 Db 534 MKKTVLTGCTGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 593
 OY 541 AVWTGYSNRLTFLVGNGLTVAAKVYRSMMTYLSEGSNPEDWNIPEGLYRNGEYFNGAR 600
 Db 594 AVWTGYSNRLTFLVGNGLTVAAKVYRSMMTYLSEGSNPEDWNIPEGLYRNGEYFNGAR 653
 OY 601 STMNSPAPQOPSTESSSSSDSTSSSTPTSTNNSTTTNNNTQOONNTTPOQONON 660
 Db 654 STMNSPAPQOPSTESSSSSDSTSSSTPTSTNNSTTTNNNTQOONNTTPOQONON 713
 OY 661 POPAP 666
 Db 714 POPAP 719

RESULT 6

O9WM11 PRELIMINARY: PRT: 719 AA.

AC O9WM11: 01-NOV-1999 (TREMREL. 12, Created)

DT 01-NOV-1999 (TREMREL. 12, Last sequence update)

DT 01-JUN-2001 (TREMREL. 17, Last annotation update)

DE PENICILLIN-BINDING PROTEIN 1A.

GN PBPIA.

OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OX NCBI_TaxID=1313;

RP SEQUENCE FROM N.A.

RC STRAIN-URU-E159, M134, URU-E135, AND URU-E157;

RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;

RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V

replacements of the cpsA-pbp1a region."

RU Microbiology 145:0-0(1999).

DR EMBL, AF138890; AAD43073.1; -

DR EMBL, AF138887; AAD43070.1; -

DR EMBL, AF138888; AAD43071.1; -

DR EMBL, AF138889; AAD43072.1; -

DR EMBL, AF159448; AAD48430.1; -

DR InterPro: IPR001264; Transglycosyl.

DR InterPro: IPR001460; Transpeptidase.

DR Pfam: PF00912; Transglycosyl; 1.

DR Pfam: PF00905; Transpeptidase; 1.

DR Prodom: PD001895; Transglycosyl; 1.

SQ SEQUENCE 719 AA; 79701 MW; 3EDC8A21C94DAC87 CRC64;

Query Match 94.7%; Score 3299; DB 2; Length 719;
 Best Local Similarity 93.4%; Pred. No. 2.4e-171;
 Matches 622; Conservative 30; Mismatches 14; Indels 0; Gaps 0;

OY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
 Db 54 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
 OY 61 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSRKAQEWLAIOLEKATKOEILTYINK 120
 Db 114 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSRKAQEWLAIOLEKATKOEILTYINK 173
 OY 121 VYMSNGNYGMQTAQNYKYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPBAADRNLVY 180


```
Db 174 VYMSGNTGMOTAAQNYGKDLNNLSLPOLALLAGMPAPNOYDPSHPEAAQDRRLVL 233
QY 181 SEMKNOGYISAQYKAVNTPTDGLQSLKASNPAYMDNTLKEVINQVEETGYNLT 240
Db 234 SEMKNOGYISAQYKAVNTPTDGLQSLKASNPAYMDNTLKEVINQVEETGYNLT 293
QY 241 TGMVYTNVDOEAQKHLMDIYNTEYVAYPDELOVASTIVDSNGKYIAOLGARHOSN 300
Db 294 TGMVYTNVDOEAQKHLMDIYNTEYVAYPDELOVASTIVDSNGKYIAOLGARHOSN 353
QY 301 VSEFGINQAVETNRDWSMKPTTDYAPALEYGVDSSTATIYHDEBYNPGTTPYNNDR 360
Db 354 VSEFGINQAVETNRDWSMKPTTDYAPALEYGVDSSTATIYHDEBYNPGTTPYNNDR 413
QY 361 GFYGNITLOVALQOOSRNPAYETTLNKVGLNRAKTFPLNGIDYPSIHSNAISSNTESD 420
Db 414 AYFGNITLOVALQOOSRNPAYETTLNKVGLNRAKTFPLNGIDYPSIHSNAISSNTESN 473
QY 421 KRYGASSEKMAAAYAAAFANGGTYRKPMYIHKVFSDSGSEKESNVGTRAMKETAYMMTD 480
Db 474 KRYGASSEKMAAAYAAAFANGGTYRKPMYIHKVFSDSGSEKESNVGTRAMKETAYMMTD 533
QY 481 MKKTYLTYGTGRNAYLAPLPAQAGKTGTSNYTDELENIHKTQSFAPADELFGYTRKYSM 540
Db 534 MKKTYLTYGTGRNAYLAPLPAQAGKTGTSNYTDELENIHKTQSFAPADELFGYTRKYSM 593
QY 541 AVMTGYSNRLTPLYNGTLVAAKYRSMNTYLSGNSPEDNNIPEGLRNGEFVFNKNGAR 600
Db 594 AVMTGYSNRLTPLYNGTLVAAKYRSMNTYLSGNSPEDNNIPEGLRNGEFVFNKNGAR 653
QY 601 STWNSPAPQOPSTRESSSSSDSSSTPSSTNNSTNNPNNNTQOOSNTTPOQONON 660
Db 654 STWNSPAPQOPSTRESSSSSDSSSTPSSTNNSTNNPNNNTQOOSNTTPOQONON 713
QY 661 POPAP 666
Db 714 POPAP 719

RESULT 7
Q9RET4 PRELIMINARY; PRT; 719 AA.
AC Q9RET4;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP 22861;
RX MEDLINE=21432820; PubMed=11549185;
RA Ferroni A., Berche P.;
RT "Alterations to penicillin-binding proteins 1a, 2b and 2x amongst
penicillin-resistant clinical isolates of Streptococcus pneumoniae
serotype 23F from the nasopharyngeal flora of children.";
RL J. Med. Microbiol. 50:828-832(2001).
DR EMBL; AF210751; AAF17261.1; -
DR InterPro; IPR001264; Transglycosyl.
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF00905; Transglycosyl. 1.
DR Pfam; PF00905; Transpeptidase.
DR PRODOM; PD001895; Transglycosyl. 1.
DR SEQUENCE 719 AA; 79567 MW; 7B6EDCCBB8CF286 CRC64;
```

Query Match 94.0%; Score 3274; DB 2: Length 719;
Best Local Similarity 92.9%; Pred. No. 5.5e-170;
Matches 619; Conservative 31; Mismatches 16; Indels 0; Gaps 0;

```
QY 1 KIYDNKNQIADLGSERRVNAQANDIPDYLKAVISIEDHREDFHGRGIDTTRILGAFLRN 60
Db 54 KIYDNKNQIADLGSERRVNAQANDIPDYLKAVISIEDHREDFHGRGIDTTRILGAFLRN 113
QY 61 LOSNLSGGSTLTQOLLKLTFFSTSDQTSRKAQEMALAIQLEQKATKQELIYYINK 120
Db 114 LOSNLSGGSTLTQOLLKLTFFSTSDQTSRKAQEMALAIQLEQKATKQELIYYINK 173
QY 121 VYMSGNTGMOTAAQNYGKDLNNLSLPOLALLAGMPAPNOYDPSHPEAAQDRRLVL 180
Db 174 VYMSGNTGMOTAAQNYGKDLNNLSLPOLALLAGMPAPNOYDPSHPEAAQDRRLVL 233
QY 181 SEMKNOGYISAQYKAVNTPTDGLQSLKASNPAYMDNTLKEVINQVEETGYNLT 240
Db 234 SEMKNOGYISAQYKAVNTPTDGLQSLKASNPAYMDNTLKEVINQVEETGYNLT 293
QY 241 TGMVYTNVDOEAQKHLMDIYNTEYVAYPDELOVASTIVDSNGKYIAOLGARHOSN 300
Db 294 TGMVYTNVDOEAQKHLMDIYNTEYVAYPDELOVASTIVDSNGKYIAOLGARHOSN 353
QY 301 VSEFGINQAVETNRDWSMKPTTDYAPALEYGVDSSTATIYHDEBYNPGTTPYNNDR 360
Db 354 VSEFGINQAVETNRDWSMKPTTDYAPALEYGVDSSTATIYHDEBYNPGTTPYNNDR 413
QY 361 GFYGNITLOVALQOOSRNPAYETTLNKVGLNRAKTFPLNGIDYPSIHSNAISSNTESD 420
Db 414 AYFGNITLOVALQOOSRNPAYETTLNKVGLNRAKTFPLNGIDYPSIHSNAISSNTESN 473
QY 421 KRYGASSEKMAAAYAAAFANGGTYRKPMYIHKVFSDSGSEKESNVGTRAMKETAYMMTD 480
Db 474 KRYGASSEKMAAAYAAAFANGGTYRKPMYIHKVFSDSGSEKESNVGTRAMKETAYMMTD 533
QY 481 MKKTYLTYGTGRNAYLAPLPAQAGKTGTSNYTDELENIHKTQSFAPADELFGYTRKYSM 540
Db 534 MKKTYLTYGTGRNAYLAPLPAQAGKTGTSNYTDELENIHKTQSFAPADELFGYTRKYSM 593
QY 541 AVMTGYSNRLTPLYNGTLVAAKYRSMNTYLSGNSPEDNNIPEGLRNGEFVFNKNGAR 600
Db 594 AVMTGYSNRLTPLYNGTLVAAKYRSMNTYLSGNSPEDNNIPEGLRNGEFVFNKNGAR 653
QY 601 STWNSPAPQOPSTRESSSSSDSSSTPSSTNNSTNNPNNNTQOOSNTTPOQONON 660
Db 654 STWNSPAPQOPSTRESSSSSDSSSTPSSTNNSTNNPNNNTQOOSNTTPOQONON 713
QY 661 POPAP 666
Db 714 POPAP 719

RESULT 8
Q57114 PRELIMINARY; PRT; 719 AA.
AC Q57114;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C5111 (CEPHALOSPORIN-RESISTANT CLINICAL ISOLATE);
RX MEDLINE=96012191; PubMed=7574521;
RA Coffey T.J., Daniels M., McDougal L.K., Dowson C.G., Tenover F.C.,
Spratt B.G.;
RT "Genetic analysis of clinical isolates of Streptococcus pneumoniae
with high-level resistance to expanded-spectrum cephalosporins.";
RL Antimicrob. Agents Chemother. 39:1306-1313(1995).
DR EMBL; Z49095; CAA88918.1; -
```


DR EMBL: Z49094; CAA88917.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
SQ SEQUENCE 719 AA; 79712 MW; 5396607C0E67D06E CRC64;

Query Match 93.2%; Score 3247; DB 2; Length 719;
Best Local Similarity 91.9%; Pred. No. 1,6e-168;
Matches 612; Conservative 33; Mismatches 21; Indels 0; Gaps 0;

OY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDITRILGAFLRN 60
DB 54 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDITRILGAFLRN 113
OY 61 LOSNSLQGGSTLTQOLIKLTYFSTSDQTSRKAQEWALAIQLEQKATKQELIYYINK 120
DB 114 LOSNSLQGGSTLTQOLIKLTYFSTSDQTSRKAQEWALAIQLEQKATKQELIYYINK 173
OY 121 VYMSNGNYGMQTAQNYKQKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRRNLV 180
DB 174 VYMSNGNYGMQTAQNYKQKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRRNLV 233
OY 181 SEMKNGYISAEQYKAVNTPITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 240
DB 234 SEMKNGYISAEQYKAVNTPITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 293
OY 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVYAYPDELOVASTIVDSNGKVIAQLGARHOSSN 300
DB 294 TGMDEVYTNVDOEAQKHLMDIYNSDYVSPDDLOVASTIVDSNGKVIAQLGARHOSSN 353
OY 301 VSEFGINQAVETNRDWSSTMKPTTDVAPALEYGYDSTATIVHDEPNYNGTNPVYNNMR 360
DB 354 VSEFGINQAVETNRDWSSTMKPTTDVAPALEYGYDSTATIVHDEPNYNGTNPVYNNMR 413
OY 361 GYFGNTITQYALQOQRNVAVETLNKVGINRAKTFPLNGIGIDYPSIHYNSNAISSNTTESD 420
DB 414 VYFGNTITQYALQOQRNVAVETLNKVGIDRAKTFPLNGIGIDYPSIHYNSNAISSNTTESN 473
OY 421 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAKETAYAMMD 480
DB 474 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAKETAYAMMD 533
OY 481 MKKTYLTYGTGRNAYLAMPLOAGKTGTSNYTDEIEIHKYIKNGYVAPDDEMFVGYTRKYAM 540
DB 534 MKKTYLTYGTGRNAYLAMPLOAGKTGTSNYTDEIEIHKYIKNGYVAPDDEMFVGYTRKYAM 593
OY 541 AVMTGYSNRLTPLVGNGLTVAAYKVRSMYTLSEGSNPEDMNIPEGLYRNGEYVFNKNGAR 600
DB 594 AVMTGYSNRLTPILIGDGLVAGKYVRSMTIYLSDDQPCDWMIPDGLYRNGEYVFNKNGAR 653
OY 601 STWNSPAPQOPSTESSSSSDSSTSQSSSTPTSTNNSTTTNPNNTQOOSNTTPOOONON 660
DB 654 STWNSPAPQOPSTESSSSSDSSTSQSSSTPTSTNNSTTTNPNNTQOOSNTTPOOONON 713
OY 661 POPAOP 666
DB 714 POPAOP 719

RESULT 9
ID Q9RET6 PRELIMINARY; PRT; 719 AA.
AC Q9RET6;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP 1053;
RX MEDLINE=21432820; PubMed=11549185;
RA Ferroul A., Berche P.;
RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst penicillin-resistant clinical isolates of Streptococcus pneumoniae";
RT serotype 23F from the nasopharyngeal flora of children.";
RL J. Med. Microbiol. 50:828-832(2001).
DR EMBL: AF210749; AAF17259.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
SQ SEQUENCE 719 AA; 79638 MW; 1DAE3C5937048E16 CRC64;

Query Match 93.2%; Score 3246; DB 2; Length 719;
Best Local Similarity 92.0%; Pred. No. 1,8e-168;
Matches 613; Conservative 35; Mismatches 18; Indels 0; Gaps 0;

OY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDITRILGAFLRN 60
DB 54 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDITRILGAFLRN 113
OY 61 LOSNSLQGGSTLTQOLIKLTYFSTSDQTSRKAQEWALAIQLEQKATKQELIYYINK 120
DB 114 LOSNSLQGGSTLTQOLIKLTYFSTSDQTSRKAQEWALAIQLEQKATKQELIYYINK 173
OY 121 VYMSNGNYGMQTAQNYKQKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRRNLV 180
DB 174 VYMSNGNYGMQTAQNYKQKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRRNLV 233
OY 181 SEMKNGYISAEQYKAVNTPITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 240
DB 234 SEMKNGYISAEQYKAVNTPITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 293
OY 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVYAYPDELOVASTIVDSNGKVIAQLGARHOSSN 300
DB 294 TGMDEVYTNVDOEAQKHLMDIYNSDYVSPDDLOVASTIVDSNGKVIAQLGARHOSSN 353
OY 301 VSEFGINQAVETNRDWSSTMKPTTDVAPALEYGYDSTATIVHDEPNYNGTNPVYNNMR 360
DB 354 VSEFGINQAVETNRDWSSTMKPTTDVAPALEYGYDSTATIVHDEPNYNGTNPVYNNMR 413
OY 361 GYFGNTITQYALQOQRNVAVETLNKVGINRAKTFPLNGIGIDYPSIHYNSNAISSNTTESD 420
DB 414 AVYFGNTITQYALQOQRNVAVETLNKVGIDRAKTFPLNGIGIDYPSIHYNSNAISSNTTESN 473
OY 421 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAKETAYAMMD 480
DB 474 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAKETAYAMMD 533
OY 481 MKKTYLTYGTGRNAYLAMPLOAGKTGTSNYTDEIEIHKYIKNGYVAPDDEMFVGYTRKYAM 540
DB 534 MKKTYLTYGTGRNAYLAMPLOAGKTGTSNYTDEIEIHKYIKNGYVAPDDEMFVGYTRKYAM 593
OY 541 AVMTGYSNRLTPLVGNGLTVAAYKVRSMYTLSEGSNPEDMNIPEGLYRNGEYVFNKNGAR 600
DB 594 AVMTGYSNRLTPILIGDGLVAGKYVRSMTIYLSDDQPCDWMIPDGLYRNGEYVFNKNGAR 653
OY 601 STWNSPAPQOPSTESSSSSDSSTSQSSSTPTSTNNSTTTNPNNTQOOSNTTPOOONON 660
DB 654 STWNSPAPQOPSTESSSSSDSSTSQSSSTPTSTNNSTTTNPNNTQOOSNTTPOOONON 713
OY 661 POPAOP 666
DB 714 POPAOP 719

SEQ	SEQUENCE	719 AA: 79750 MM: 489B6C6EC0CB861CD CRC64:
054949	PRELIMINARY;	PRT; 719 AA.
054949		
054949		
01-NOV-1996	(TREMBLrel. 01, Created)	
01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).		
PONA.		
Streptococcus pneumoniae.		
Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae.		
Streptococcus.		
NCBI_TaxID=1313;		
(1)		
SEQUENCE FROM N.A.		
SPRAIN=56742;		
MEDLINE=93010977; PubMed=1396576;		
Martin C., Sibold C., Hakenbeck R.;		
"Relationship of penicillin-binding protein 1a genes from different		
clones of penicillin-resistant Streptococcus pneumoniae isolated in		
South Africa and Spain."		
EMBO J. 11:3831-3836(1992).		
EMBL; X67869; CAA48069.1; -		
InterPro; IPR001264; Transglycosyl.		
InterPro; IPR001460; Transpeptidase.		
Pfam; PF009912; Transglycosyl; 1.		
Pfam; PF009905; Transpeptidase; 1.		
ProDom; PD001895; Transglycosyl; 1.		
NON_TER	719	719

Query Match	93.28	Score 3246	DB 2	Length 719
Best Local Similarity	92.08	Pred. No. 1.8e-168		
Matches 613, Conservative	31	Mismatches 22	Indels 0	Gaps 0

Qy	1	KTIIDNKQNLJNDLGSERVYNAQANDIPDULYKALVSIEDHRFPBHRKIDITRILIGAFLRN	60
Db	54	KIYDNRKQNLJNDLGSERVYNAQANDIPDULYKALVSIEDHRFPBHRKIDIRIGAFLRN	113
Qy	61	LOSNSLOGSFTLTOOLIKLTYFTSTSQOTSRAQOEAMLAIDQEKATQOEILTYINK	120
Db	114	LOSNSLOGSFTLTOOLIKLTYFTSTSDQOTSRAQOEAMLAIDQEKATQOEILTYINK	173
Qy	121	YVMSNGNYGMQTAANYKQDNLNLSPOLALLAGMPQAPNQDYPYSHPEAADOBRNLVL	180
Db	174	YVMSNGNYGMQTAANYKQDNLNLSPOLALLAGMPQAPNQDYPYSHPEAADOBRNLVL	233
Qy	181	SEMKNOGYSIAEYQEKAVNTPITDGLQSLKSASNYPAYMONYKEVINQOEETGYNLLT	240
Db	234	SEMKNOGYSIAEYQEKAVNTPITDGLQSLKSASNYPAYMONYKEVINQOEETGYNLLT	293
Qy	241	TGMQVYTNVQDOEAKHLMIDITNDBEYAAYPDDELQVASTIVDVSNGKVIQOLGARHOSN	300
Db	294	TGMQVYTNVQDOEAKHLMIDITNSQVYSPDDELQVASTIVDVSNGKVIQOLGARHOSN	353
Qy	301	VSEFGLNOAVETNRQMGSTMKPITDYADALEGYVDSTATTYHDEPNYPCGNTNPVYMMDR	360
Db	354	VSEFGLNOAVETNRQMGSMKPIITDYAALIEGYVDSTATTYVNDIPYPCGTSPIYVMMDR	413
Qy	361	GYFGNITLQVALQOOSRNVPAVETLNKVGLENRAKTFGLNGLGDVPSIYSHNAISNTESD	420
Db	414	AYFGNITLQVALQOOSRNVTAETLNKVGLEDBRAKTFGLNGLGDVPSIYSHNAISNTESN	473
Qy	421	KKYGASSSEKMAAAVAAFPANGSTYKKPMYIHKVYFSDSESEKFEFNVGRAKETTAYMTD	480
Db	474	KOYGASSSEKMAAAVAAFGANGTYKKPMYIHKVYFSDSKKEFNSVGRAKETTAYMMTD	533
Qy	481	MMKTYLVTGTGRNRYLWMLPAGTGTGNSNTDEIEKHITKSOPVADDELFCAGYTRYSM	540
Db	534	MMKTYLVTGTGRGAYLPRMLPAGKTGTSNITDEVEVHNKIKTGIVADEMFVGTTRYYSM	593
Qy	541	AVWGTYSNRLPLVNGNLJVAAKVYRSMMTYLSGCSNPBDMNITDEGLYRNGEYFVKKGAR	600

[illegible]

RESULT	11		
054947			
ID	054947	PRELIMINARY;	PRT; 719 AA.
AC	054947;		
DT	01-NOV-1996 (TREMBlrel. 01, Created)		
DT	01-NOV-1996 (TREMBlrel. 01, Last sequence update)		
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)		
DE	PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).		
GN	PONA.		
OS	Streptococcus pneumoniae.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;		
OX	Streptococcus.		
OX	NCBI_TaxID=1313;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=681;		
RX	MEDLINE=93010977; PubMed=1396576;		
RA	Martin C., Sibold C., Hakenbeck R.;		
RT	"Relatedness of penicillin-binding protein 1a genes from different		
RT	clones of penicillin-resistant Streptococcus pneumoniae isolated in		
RT	South Africa and Spain.";		
RL	EMBO J. 11:3831-3836(1992).		
DR	EMBL; X67866; CAA4806.1; -		
DR	InterPro; IPR001264; Transglycosyl.		
DR	InterPro; IPR001460; Transpeptidase.		
DR	Pfam; PF00912; Transglycosyl. 1.		
DR	Pfam; PF00905; Transpeptidase; 1.		
DR	ProDom; PD001895; Transglycosyl. 1.		
FT	NON_TER	719	719
SO	SEQUENCE	719 AA;	79710 MW; D90EB42D5380D9E1 CRC64;

Query Match	93.18;	Score 3245;	DB 2;	Length 719;
Best Local Similarity	92.08;	Pred. No. 2.1e-168;		
Matches 613; Conservative	35;	Mismatches 18;	Indels 0;	Gaps 0;

Qy	1	KIIDNKNOIADLGSERVVNAQANDIPDDLYKALVSIEDHRRFPHRSDITRIILGAFLRN	60
Db	54	KIYDKNOLLIADLGSERVVNAQANDIPDDLYKALVSIEDHRRFPHRSDISIRILGAFLRN	113
Qy	61	LOSNSLOGGSTLTPOOLIKLFEVSTSTSPQTSKRAQEWMLAIOLEOKATKOELLTYINK	120
Db	114	LOSNSLOGGSTLTPOOLIKLFTSTSTSPQTSKRAQEWMLAIOLEOKATKOELLTYINK	173
Qy	121	YVMSGNGNGMOTAAQNYNGKDLNMLSPOLALLAGMPOAENQYDPYSHPEAADRRNLVYL	180
Db	174	YVMSGNGNGMOTAAQNYNGKDLNMLSPOLALLAGMPOAENQYDPYSHPEAADRRNLVYL	233
Qy	181	SEMKNOGYISAEQYKAVNPIETDGLQSLKSASNNPAYMNNYLKEVINQVEETGYNLLT	240
Db	234	SEMKNOGYISAEQYKAVNPIETDGLQSLKSASNNPAYMNNYLKEVINQVEETGYNLLT	293
Qy	241	TGMDVYTTWVDEDAQKHLMDIYNDEYVAYYPDDELQVASTYVDVSNKGVIHQAGRHSSN	300
Db	294	TGMDVYTTWVDEDAQKHLMDIYNQYVYSPDDELQVASTVVDVSNKGVIHQAGRHQASN	353
Qy	301	VSFQGNQAVENNRBMGSMKRTITTYAPALEEGYVDSPATTVHBEPPYVPCETNNPVYNNMWR	360
Db	354	VSFQGNQAVENNRBMGSMKRTITTYAPALEEGYVDSPATTVNNDDIPNTPCTSPVPVNNMWR	413
Qy	361	GYFGNITLTQVYALQOOSRNVPAAETLNKGLNANKAFLLNGLQIDVPSIHYSAIISNTTESD	420


```

|||||
414 AFGNITLQYALQOSRNVAVETLNKVGIDRAKTFNLNGIGIDYPSMHYANAISSNTTESN 473
QY 421 KYGASSEKMAAAYAFANGGYTKPMYIHKVFSGSKSEKFSNVGTRAKKETATAMMD 480
DB 474 KÖGASSEKMAAAYAFANGGYIHKPMYIHKVFSGSKSEKFSNVGTRAKKETATAMME 533
QY 481 MKTIVLTGTRNAYLAWLPQAGKTGTSNYTDEIEINHITQSOFVAPDELFAGYTRKYSM 540
DB 534 MKKTVLAVGTGRCAYLPLWLPQAGKTGTSNYTDEIEIKHISTGYVAPDEMFVYTRKYSM 593
QY 541 AWTGYSNRLTPLVGNGLVYAAKVYRSMYIYSEGSNPEDWNIPEGLYRNGEFVFKNGAR 600
DB 594 AWTGYSNRLTPLVGDGFLVYAAKVYRSMITLYSEDTHPEDWMPDGLFRNGEFVFKNGAR 653
QY 601 STWNSPAPQOPSTESSSSSDSSTQSSSTPTSTNNSTTTNNNTQOOSNTTPOOON 660
DB 654 STWNSPAPQOPSTESSSSSDSSTQSSSTPTSTNNSTTTNNNTQOOSNTTPOOON 713
QY 661 POPAP 666
DB 714 POPAP 719

RESULT 12
Q54946 PRELIMINARY: PRT: 719 AA.
AC Q54946:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA OR PBPLA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=113;
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=456;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Stibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain.";
RL EMO J. 11:3831-3836(1992).
RN [2]
RP SEQUENCE OF 310-619 FROM N.A.
RC STRAIN=8303, AND 35193;
RX Smith A.M., Klugman K.P.;
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 320-717 FROM N.A.
RC STRAIN=417/246, #20/B98, AND #27/SHA3;
RX MEDLINE=98409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a thr-371 substitution in a conserved amino acid motif
RT of penicillin-binding protein 1A with penicillin resistance of
RT Streptococcus pneumoniae.";
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=SP-665;
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
RT clone of Streptococcus pneumoniae arose by large recombinational
RT replacements of the cpsA-pbpA region.";
RL Microbiology 0:0-0(1999).
DR EMBL: X67868; CAA48068.1; -
DR EMBL: AF046230; AAC24695.1; -
DR EMBL: AB006876; BAA32072.1; -
DR EMBL: AF139883; AAD43066.1; -
DR InterPro: IPR001264; Transglycosyl.

```

```

DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl. 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR Prodom: PD001895; Transglycosyl. 1.
FT NON_TER 719 719
SQ SEQUENCE 719 AA; 79736 MM; F090F9152C834DC CRC64;

Query Match 93.1%; Score 3243; DB 2; Length 719;
Best local Similarity 91.9%; Pred. No. 2,7e-168;
Matches 612; Conservative 32; Mismatches 22; Indels 0; Gaps 0;

QY 1 KIYDNKNQLIADIGSSRRVNAQANDIPTDLVKAIVSIEDHFFDHRGIDITRILGAFLRN 60
DB 54 KIYDNKNQLIADIGSSRRVNAQANDIPTDLVKAIVSIEDHFFDHRGIDISIRILGAFLRN 113
QY 61 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKRKADEMTALIOLEKATKQELITLYINK 120
DB 114 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKRKADEMTALIOLEKATKQELITLYINK 173
QY 121 VYKSGNGYGMQTAQNYVYKDLNNLSLPOLALLAGHPQAPNOVDYPSHPAADDNRNLV 180
DB 174 VYKSGNGYGMQTAQNYVYKDLNNLSLPOLALLAGHPQAPNOVDYPSHPAADDNRNLV 233
QY 181 SEMKNOGYISAEQYERAVNTPITDGLQSLKSASNYPAYMDNYLKEYI NOVEEETGYNLTL 240
DB 234 SEMKNOGYISAEQYERAVNTPITDGLQSLKSASNYPAYMDNYLKEYI NOVEEETGYNLTL 293
QY 241 TGMADVNTNDQEAQKHLMDIYNTDEVYVAPDELOVASTIVDSNGKVIAQLGARHOSSN 300
DB 294 TGMADVNTNDQEAQKHLMDIYNSDQYVSPDDLOVASTIVDSNGKVIAQLGARHOASN 353
QY 301 VSFSGINQAVETNRDQSTKPIITDYVPALEYGYDSTATIVHDEPNYPCGTNPPVNMDR 360
DB 354 VSFSGINQAVETNRDQSTKPIITDYVPALEYGYDSTATIVHDEPNYPCGTNPPVNMDR 413
QY 361 GYFGNITLQYALQOSRNVAVETLNKVGIDRAKTFNLNGIGIDYPSIHSNAISSNTTESD 420
DB 414 AFGNITLQYALQOSRNVAVETLNKVGIDRAKTFNLNGIGIDYPSMHYANAISSNTTESN 473
QY 421 KYGASSEKMAAAYAFANGGYTKPMYIHKVFSGSKSEKFSNVGTRAKKETATAMMD 480
DB 474 KÖGASSEKMAAAYAFANGGYIHKPMYIHKVFSGSKSEKFSNVGTRAKKETATAMMD 533
QY 481 MKTIVLTGTRNAYLAWLPQAGKTGTSNYTDEIEINHITQSOFVAPDELFAGYTRKYSM 540
DB 534 MKTIVLTGTRNAYLAWLPQAGKTGTSNYTDEIEINHITQSOFVAPDELFAGYTRKYSM 593
QY 541 AWTGYSNRLTPLVGNGLVYAAKVYRSMYIYSEGSNPEDWNIPEGLYRNGEFVFKNGAR 600
DB 594 AWTGYSNRLTPLVGDGFLVYAAKVYRSMITLYSEDTHPEDWMPDGLFRNGEFVFKNGAR 653
QY 601 STWNSPAPQOPSTESSSSSDSSTQSSSTPTSTNNSTTTNNNTQOOSNTTPOOON 660
DB 654 STWNSPAPQOPSTESSSSSDSSTQSSSTPTSTNNSTTTNNNTQOOSNTTPOOON 713
QY 661 POPAP 666
DB 714 POPAP 719

RESULT 13
Q54948 PRELIMINARY: PRT: 719 AA.
AC Q54948:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.

```


NCBI_TaxID=1313;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=670;
 RX MEDLINE=93010977; PubMed=1396576;
 RA Matlin C., Sibold C., Hakenbeck R.;
 RT "Relatedness of penicillin-binding protein 1a genes from different
 ST clones of penicillin-resistant Streptococcus pneumoniae isolated in
 RL South Africa and Spain."
 DR EMBL: J. 11:3831-3836(1992).
 DR EMBL: X67867; CA48067.1; -;
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 FT NON_TER 719 719
 SQ SEQUENCE 719 AA; 79611 MW; 1DBFDDA93704930B CRC64;

Query Match 93.0%; Score 3241; DB 2; Length 719;
 Best Local Similarity 91.9%; Pred. No. 3.4e-168;
 Matches 612; Conservative 36; Mismatches 18; Indels 0; Gaps 0;

OY 1 KIYNKNQNLADLGSERRVNAQANDIPTDLVKAIVSIEDHREFPDHRCIDITRIIGAFARN 60
 DB 54 KIYNKNQNLADLGSERRVNAQANDIPTDLVKAIVSIEDHREFPDHRCIDITRIIGAFARN 113
 OY 61 LOSNSLOGGSTLTQOLIKLTYFSTSTSDQTSRKAQEMALAIQLEOKATKOEILTYINK 120
 DB 114 LOSNSLOGGSTLTQOLIKLTYFSTSTSDQTSRKAQEMALAIQLEOKATKOEILTYINK 173
 OY 121 VYMSNGNGMOTAAQNYGKDLNLSLPOLALLAGMPAPQOYPSHPEAODRNLVL 180
 DB 174 VYMSNGNGMOTAAQNYGKDLNLSLPOLALLAGMPAPQOYPSHPEAODRNLVL 233
 OY 181 SEMNNOGYISAEOYEKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYNLT 240
 DB 234 SEMNNOGYISAEOYEKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYNLT 293
 OY 241 TGMVYTNVDOEAKHLMIDYNTDEYVAYPDDELQVASTIVDVNSGKVIQALGARHOSN 300
 DB 294 TGMVYTNVDOEAKHLMIDYNTDEYVAYPDDELQVASTIVDVNSGKVIQALGARHOSN 353
 OY 301 VSEFGINAVETNRDGMGSMKPIIDYAPALEGYVDSTATYHDEPYNPGNTPYVNMDR 360
 DB 354 VSEFGINAVETNRDGMGSMKPIIDYAPALEGYVDSTATYHDEPYNPGNTPYVNMDR 413
 OY 361 GYFGNITLQVALQOOSRNVPAYETLNKVLNRAKTFNLGIDYPSIHYSNAISSNTTESD 420
 DB 414 AYGFNITLQVALQOOSRNVPAYETLNKVLNRAKTFNLGIDYPSIHYSNAISSNTTESD 473
 OY 421 KKYGASSSEKMAAAYAAAFANGSTYRKPMYIHKVYVSDGSEKFEFSVNGTRAMEETAYMMD 480
 DB 474 KKYGASSSEKMAAAYAAAFANGSTYRKPMYIHKVYVSDGSEKFEFSVNGTRAMEETAYMMD 533
 OY 481 MKKTVLTYGTRNAYLWLPQAGKTGTSNYTDEELENHRIKTSQVAPDELFGAGTRKXSM 540
 DB 534 MKKTVLTYGTRNAYLWLPQAGKTGTSNYTDEELENHRIKTSQVAPDELFGAGTRKXSM 593
 OY 541 AVMTGYSNRLTPLYNGLTVAAYKYSMTYLSGSPNEDNINIEGGLYRNGEYFKNGAR 600
 DB 594 AVMTGYSNRLTPLYNGLTVAAYKYSMTYLSGSPNEDNINIEGGLYRNGEYFKNGAR 653
 OY 601 STWNSPAPQPPSTRESSSSSDSTSSSTTSTNNSTTNNSTTNNSTTNNSTTNNSTTNNSTTNN 660
 DB 654 STWNSPAPQPPSTRESSSSSDSTSSSTTSTNNSTTNNSTTNNSTTNNSTTNNSTTNNSTTNN 713
 OY 661 POPAOP 666
 DB 714 POPAOP 719

RESULT 14
 O9RET5
 ID O9RET5 PRELIMINARY; PRT: 719 AA.
 AC O9RET5;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBP1A.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RP SEQUENCE FROM N.A.
 RC STRAIN=SP 1470;
 RX MEDLINE=21432820; PubMed=11549185;
 RA Ferroni A., Berche P.;
 RT "Alterations to penicillin-binding proteins 1a, 2b and 2x amongst
 ST penicillin-resistant clinical isolates of Streptococcus pneumoniae
 RT serotype 23F from the nasopharyngeal flora of children."
 RL J. Med. Microbiol. 50:828-832(2001).
 DR EMBL: AF210750; AAF17260.1; -;
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 FT NON_TER 719 719
 SQ SEQUENCE 719 AA; 79734 MW; 3BE77571DBC86496 CRC64;

Query Match 92.7%; Score 3231; DB 2; Length 719;
 Best Local Similarity 91.4%; Pred. No. 1.2e-167;
 Matches 609; Conservative 34; Mismatches 23; Indels 0; Gaps 0;

OY 1 KIYNKNQNLADLGSERRVNAQANDIPTDLVKAIVSIEDHREFPDHRCIDITRIIGAFARN 60
 DB 54 KIYNKNQNLADLGSERRVNAQANDIPTDLVKAIVSIEDHREFPDHRCIDITRIIGAFARN 113
 OY 61 LOSNSLOGGSTLTQOLIKLTYFSTSTSDQTSRKAQEMALAIQLEOKATKOEILTYINK 120
 DB 114 LOSNSLOGGSTLTQOLIKLTYFSTSTSDQTSRKAQEMALAIQLEOKATKOEILTYINK 173
 OY 121 VYMSNGNGMOTAAQNYGKDLNLSLPOLALLAGMPAPQOYPSHPEAODRNLVL 180
 DB 174 VYMSNGNGMOTAAQNYGKDLNLSLPOLALLAGMPAPQOYPSHPEAODRNLVL 233
 OY 181 SEMNNOGYISAEOYEKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYNLT 240
 DB 234 SEMNNOGYISAEOYEKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYNLT 293
 OY 241 TGMVYTNVDOEAKHLMIDYNTDEYVAYPDDELQVASTIVDVNSGKVIQALGARHOSN 300
 DB 294 TGMVYTNVDOEAKHLMIDYNTDEYVAYPDDELQVASTIVDVNSGKVIQALGARHOSN 353
 OY 301 VSEFGINAVETNRDGMGSMKPIIDYAPALEGYVDSTATYHDEPYNPGNTPYVNMDR 360
 DB 354 VSEFGINAVETNRDGMGSMKPIIDYAPALEGYVDSTATYHDEPYNPGNTPYVNMDR 413
 OY 361 GYFGNITLQVALQOOSRNVPAYETLNKVLNRAKTFNLGIDYPSIHYSNAISSNTTESD 420
 DB 414 AYGFNITLQVALQOOSRNVPAYETLNKVLNRAKTFNLGIDYPSIHYSNAISSNTTESD 473
 OY 421 KKYGASSSEKMAAAYAAAFANGSTYRKPMYIHKVYVSDGSEKFEFSVNGTRAMEETAYMMD 480
 DB 474 KKYGASSSEKMAAAYAAAFANGSTYRKPMYIHKVYVSDGSEKFEFSVNGTRAMEETAYMMD 533
 OY 481 MKKTVLTYGTRNAYLWLPQAGKTGTSNYTDEELENHRIKTSQVAPDELFGAGTRKXSM 540
 DB 534 MKKTVLTYGTRNAYLWLPQAGKTGTSNYTDEELENHRIKTSQVAPDELFGAGTRKXSM 593
 OY 541 AVMTGYSNRLTPLYNGLTVAAYKYSMTYLSGSPNEDNINIEGGLYRNGEYFKNGAR 600
 DB 594 AVMTGYSNRLTPLYNGLTVAAYKYSMTYLSGSPNEDNINIEGGLYRNGEYFKNGAR 653

[illegible][illegible]

Db	414	YIFGNTTQYALQOOSRNTAVETLINKVGLDRAKFFLNGLGIDIPESMAYANAISNTTESN	473
Qy	421	KKYGASSERKMAAAVAEANGSTYKKPMYIHKVFE.SDGESEKEFSNVGTRAKETTTAYMTD	480
Db	474	KOYGASSERKMAAAVAEANGSTYKKPMYINKIVFSDGESEKEFPDAGTRAKETTTAYMTE	533
Qy	481	MKKTYLTGTGNNYLAALPDAGTGTNSNYDEIEENHKITSOPVADDELFCAGYTRKYS	540
Db	534	MKKTYLAAGIGRGAYLPLPLPDAGTGTINSNTYDDEIEKIKMTGYVADDEMFVGTIRKYS	593
Qy	541	AVMTGYSNRLLTPLYVNGILTVAAKYRSMYTLSEGSNPEDMNIPDGLYRNGEIFYKNGAR	600
Db	594	AVMTGYSNRLLTPLYVODGLTVAAKYRSMYTLSEGSNPEDMNIPDGLYRNGEIFYKNGAR	653
Qy	601	STWNSPAAQOPSPSTFSSSSSSDSSSTFSTNNSTTTNNPNNTQQSNTTPDOONON	660
Db	654	STWNSPAAQOPSPSTFSSSSSSDSSSTFSTNNSTTTTDPNNNTQQSNTTPDOONON	713
Qy	661	POPAPQ 666	
Db	714	POPAPQ 719	

Search completed: June 13, 2002, 08:45:21
Job time: 252 sec

THIS PAGE BLANK (USPTO)


```

Db      1 KIYDNKNOLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLEFN 60
OY      61 LOSNSLOGSSTLTLOOLIKLTFTSTSDOTISRKAOEAMLAIOLEKATKOELLTYINK 120
        |||||||
Db      61 LOSNSLOGSSTLTLOOLIKLTFTSTSDOTISRKAOEAMLAIOLEKATKOELLTYINK 120
OY      121 VYMSNGNGMOTPAQONYKGDNLNLSLPOLALLAGMPQAPNOYDPYSHEEAAODRRNLVL 180
        |||||||
Db      121 VYMSNGNGMOTPAQONYKGDNLNLSLPOLALLAGMPQAPNOYDPYSHEEAAODRRNLVL 180
OY      121 VYMSNGNGMOTPAQONYKGDNLNLSLPOLALLAGMPQAPNOYDPYSHEEAAODRRNLVL 180
        |||||||
Db      121 VYMSNGNGMOTPAQONYKGDNLNLSLPOLALLAGMPQAPNOYDPYSHEEAAODRRNLVL 180
OY      181 SEMKNOGYISAQOYEKAVNTPITDGLQSLKSASNYPAYMDNYLKEVINOVEETGYNLLT 240
        |||||||
Db      181 SEMKNOGYISAQOYEKAVNTPITDGLQSLKSASNYPAYMDNYLKEVINOVEETGYNLLT 240
OY      241 TGMADVYTNDOEAQKHLMDIYNTDEVAYAPDELOQVASTIVDSNGKVAIOGARRHOSN 300
        |||||||
Db      241 TGMADVYTNDOEAQKHLMDIYNTDEVAYAPDELOQVASTIVDSNGKVAIOGARRHOSN 300
OY      301 VSEFGINQAVETNRDMGSTMKPITDYAPALEYGYDSTATIVHDEPNYNGTNPVYNMNR 360
        |||||||
Db      301 VSEFGINQAVETNRDMGSTMKPITDYAPALEYGYDSTATIVHDEPNYNGTNPVYNMNR 360
OY      361 GYFGNITLQYALQOQSRNVPVAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNALSSNTTESD 420
        |||||||
Db      361 GYFGNITLQYALQOQSRNVPVAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNALSSNTTESD 420
OY      421 KRYGASSEKMAAAYAAAFANGGTYRPMYIHKYVSDGSEKESNVGTRAMKETTAYAMND 480
        |||||||
Db      421 KRYGASSEKMAAAYAAAFANGGTYRPMYIHKYVSDGSEKESNVGTRAMKETTAYAMND 480
OY      481 MKKTYLVYTGGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSOFVAPDELFAGYTRKYSM 540
        |||||||
Db      481 MKKTYLVYTGGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSOFVAPDELFAGYTRKYSM 540
OY      541 AWMTGYSNRLTPLYVGNGLTVAAYKVRSMNTYLSSEGSNPEDMNIPEGLYRNGEEVFNKGAR 600
        |||||||
Db      541 AWMTGYSNRLTPLYVGNGLTVAAYKVRSMNTYLSSEGSNPEDMNIPEGLYRNGEEVFNKGAR 600
OY      601 STWNSPAPQOPPESTESSSSSDSTSSQSSSTPTSTNNSTTTNNNTTQOOSNTTPOQONON 660
        |||||||
Db      601 STWNSPAPQOPPESTESSSSSDSTSSQSSSTPTSTNNSTTTNNNTTQOOSNTTPOQONON 660
OY      661 POPAOP 666
        |||||||
Db      661 POPAOP 666

```

```

APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Steiner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 354-8113
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-435-6

Query Match          99.5%; Score 3466; DB 3; Length 682;
Best Local Similarity 99.4%; Pred No. 5 4e-254;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY      1 KIYDNKNOLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLEFN 60
        |||||||
Db      17 KIYDNKNOLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLEFN 76
        |||||||
OY      61 LOSNSLOGSSTLTLOOLIKLTFTSTSDOTISRKAOEAMLAIOLEKATKOELLTYINK 120
        |||||||
Db      77 LOSNSLOGSSTLTLOOLIKLTFTSTSDOTISRKAOEAMLAIOLEKATKOELLTYINK 136
        |||||||
OY      121 VYMSNGNGMOTPAQONYKGDNLNLSLPOLALLAGMPQAPNOYDPYSHEEAAODRRNLVL 180
        |||||||
Db      137 VYMSNGNGMOTPAQONYKGDNLNLSLPOLALLAGMPQAPNOYDPYSHEEAAODRRNLVL 196
        |||||||
OY      181 SEMKNOGYISAQOYEKAVNTPITDGLQSLKSASNYPAYMDNYLKEVINOVEETGYNLLT 240
        |||||||
Db      197 SEMKNOGYISAQOYEKAVNTPITDGLQSLKSASNYPAYMDNYLKEVINOVEETGYNLLT 256
        |||||||
OY      241 TGMADVYTNDOEAQKHLMDIYNTDEVAYAPDELOQVASTIVDSNGKVAIOGARRHOSN 300
        |||||||
Db      257 TGMADVYTNDOEAQKHLMDIYNTDEVAYAPDELOQVASTIVDSNGKVAIOGARRHOSN 316
        |||||||
OY      301 VSEFGINQAVETNRDMGSTMKPITDYAPALEYGYDSTATIVHDEPNYNGTNPVYNMNR 360
        |||||||
Db      317 VSEFGINQAVETNRDMGSTMKPITDYAPALEYGYESTATIVHDEPNYNGTNPVYNMNR 376
        |||||||
OY      361 GYFGNITLQYALQOQSRNVPVAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNALSSNTTESD 420
        |||||||
Db      377 GYFGNITLQYALQOQSRNVPVAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNALSSNTTESD 436
        |||||||
OY      421 KRYGASSEKMAAAYAAAFANGGTYRPMYIHKYVSDGSEKESNVGTRAMKETTAYAMND 480
        |||||||
Db      437 KRYGASSEKMAAAYAAAFANGGTYRPMYIHKYVSDGSEKESNVGTRAMKETTAYAMND 496
        |||||||
OY      481 MKKTYLVYTGGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSOFVAPDELFAGYTRKYSM 540
        |||||||
Db      497 MKKTYLVYTGGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSOFVAPDELFAGYTRKYSM 556
        |||||||
OY      541 AWMTGYSNRLTPLYVGNGLTVAAYKVRSMNTYLSSEGSNPEDMNIPEGLYRNGEEVFNKGAR 600
        |||||||
Db      557 AWMTGYSNRLTPLYVGNGLTVAAYKVRSMNTYLSSEGSNPEDMNIPEGLYRNGEEVFNKGAR 616
        |||||||
OY      601 STWNSPAPQOPPESTESSSSSDSTSSQSSSTPTSTNNSTTTNNNTTQOOSNTTPOQONON 660
        |||||||
Db      617 STWNSPAPQOPPESTESSSSSDSTSSQSSSTPTSTNNSTTTNNNTTQOOSNTTPOQONON 676
        |||||||
OY      661 POPAOP 666
        |||||||
Db      677 POPAOP 682

```



```
RESULT 3
US-08-245-511-4
; Sequence 4, Application US/08245511
; Patent No. 5928900
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THERON
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-245-511-4

Query Match          46.5%; Score 1621; DB 2; Length 320;
Best Local Similarity 99.7%; Pred. No. 6.7e-115;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
DB 305 VSEGINQAVETNRDNG 320
|||||
RESULT 4
US-08-600-993A-4
; Sequence 4, Application US/0860093A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THERON
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-600-993A-4

Query Match          46.5%; Score 1621; DB 2; Length 320;
Best Local Similarity 99.7%; Pred. No. 6.7e-115;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```


Db 185 SEMKNOGYISABOYEKAVNTPITDGLQSLKASANTPAYMDNTLKEYINOVEERTGYNLIT 244
 QY 241 TGMDEVYTNDOEAKHLMIDYNTDEVAYPDELQVASTIVDSNGKYIAOLGARHOSSN 300
 Db 245 TGMDEVYTNDOEAKHLMIDYNTDEVAYPDELQVASTIVDSNGKYIAOLGARHOSSN 304
 QY 301 VSFGINQAVETNRDMG 316
 Db 305 VSFGINQAVETNRDMG 320

RESULT 5
 US-08-731-716-2
 ; Sequence 2, Application US/08731716
 ; Patent No. 5789202

GENERAL INFORMATION:
 APPLICANT: Hoskins, John
 APPLICANT: Jaskunas, S. Richard
 APPLICANT: Rokey, Pamela K.
 APPLICANT: Zhao, Genshi
 APPLICANT: Rostek, Paul R. Jr.
 APPLICANT: No. 5789202, Franklin H.
 TITLE OF INVENTION: Penicillin Binding Protein From
 Streptococcus Pneumoniae
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Eli Lilly and Company
 STREET: Lilly Corporate Center
 CITY: Indianapolis
 STATE: Indiana
 COUNTRY: U.S.
 ZIP: 46285
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/731,716
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Webster, Thomas D.
 REGISTRATION NUMBER: 39, 872
 REFERENCE/DOCKET NUMBER: X-10,887
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 317-276-3334
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 731 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-731-716-2

Query Match 18.3%; Score 636; DB 1; Length 731;
 Best Local Similarity 27.7%; Pred. No. 8.2e-40;
 Matches 192; Conservative 108; Mismatches 252; Indels 142; Gaps 19;

QY 2 IYDNKNQIADIGSERRNAQANDIPTDLVAKAIVSIEDHRFEDHGIDITRILGAFRLNL 61
 Db 95 IDPREKEKAGALSGQKVELTLDISKNLONNAVATEDRSFYKNDGINVGRF--FLAIV 151
 QY 62 QSNLSLGGSTLQOLIKLYFSTSDQISRAQAMLAIQEQKATQOELITYINKY 121
 Db 152 TAGRSGGSTLQOLAKNYLS--ODQYVERAKAEFFALSKRYSEQLITMPLNMA 208
 QY 122 YMSNGYMGQTAQNYGDLNLNLSPOLLAGMPOAPNOYDPSGHPPEAODRNLVLS 181
 Db 209 YFGNGWGVEDASKEYFGSASAEVSLDQAATLAGMLKGPFLNPLNSVEDSTRRIIVLO 268
 QY 182 EMKNOGYISABOYEKAVNTPIT---DGLQSLKASANTPAYMDNTLKEYI--NOVEERT 234

Db 269 NNVAAGYIDKNOETEAQVDMKTSQLDHKEYGKISDRYRSTYPAVAVNEAVSKYNLTTEE- 327
 QY 235 GYNLLTGMDEVYTNDOEAKHLMIDY-NTDEVAYPDEL-QVASTIVDSNGKYIAOL 292
 Db 328 ---IYVNGYRITELDONQANMOIYENTSLFPRAEDGTFFQSSGVVALEPTGVRGVY 384
 QY 293 GARHOSSNVSF-GINQAVETNRDMGSTMKPITDYAPALEGYGYDSTATTIVHDEPY----- 346
 Db 385 GOVADNDKDTGFNFNFNATQSKRSPGSTIKPLVYTPAVAGMALNQLDNHMQYDSYRV 444
 QY 347 -NYPGTNT---PYVNMWDGFGYGNITLQYALQOOSRVPAVETLNKGLNRKAKFTLNGLGI 401
 Db 445 DNAGIKTSREVPYIO-----SLAESLNDPAVATVNDGVDKA----- 482
 QY 402 DYPSIHYNSNAISSNTTESDKKYGASSEK-----MAAAYAFANGGT 442
 Db 483 -----FEAGEKFGLNMEKXDRVLGVALGSGVETNPLOMQAVAFANEGL 527
 QY 443 YKPKMTIHKVPSDGE-KEFSNVGTRAKETTYAMMTDMKTVLTGGRNAYLAMLPQ 501
 Db 528 MPEAHFISHIENASGOVIAHSHKNSQKRVIDKSVADKMTSMILCTFTNGTGISSSPADYVA 587
 QY 502 AGKTGTSNTDEIEHHTSQFVAPDELFCAGYTRKYSMAVWTGY-----SNRLTFLVCG 556
 Db 588 AGKTGTT---EAVFNPEYTS-----DQVYIGTTPDVPVISHWLGFTTDBNHTLASTSN 638
 QY 557 GLTVAKYRSMYTYLSEGSNPEDMNIPGLYRNGEFVRKNGARSTWNSPAPQPPSTES 616
 Db 639 G---AAHVERNIANTIL-----PYTGG 657
 QY 617 SSSSSDSTSGSSSTPSTNNSTTNPNNNTQOS 650
 Db 658 STFTVENAYKONGIAPANTKRQYQTDNDSQTDON 691

RESULT 6
 US-08-481-435-2
 ; Sequence 2, Application US/08481435
 ; Patent No. 6027906
 GENERAL INFORMATION:
 APPLICANT: Balganes, Tanjore S
 APPLICANT: Town, Christine
 TITLE OF INVENTION: No. 6027906, Polypeptides
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: White & Case
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States
 ZIP: 10036-2787
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/481,435
 FILING DATE: 10-JUL-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: IN 580/MAS/94
 FILING DATE: 01-JUL-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9404072-2
 FILING DATE: 24-NOV-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Steiner, Richard J.
 REGISTRATION NUMBER: 35,372
 REFERENCE/DOCKET NUMBER: 1103326-151
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 819-8783


```

: TELEFAX: (212) 354-8
: INFORMATION FOR SEQ ID NO
: SEQUENCE CHARACTERISTIC
: LENGTH: 828 amino acid
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
US-08-481-435-2

```

Query Match	16.3%	Score 567;	DB 3;	Length 828;
Best Local Similarity	23.4%	Pred. No. 1.7e-34;		
Matches 202;	Conservative 129;	Mismatches 244;	Indels 290;	Gaps 30;

OY	1	KIYONKQULLADJGSESRVNAQANDIPTDLKAVIVSEJEDRFPDHRIDITRIL-GAFIR	59
Db	27	QIYABODELLAIOGKERKIRIVITLDQIPREPVAKFIANEDSRFEYEHNVDPVGLFRASVA	86
OY	60	NLOSNIJGSSSTYLTOOLIKLIFYSTSTSDQTSKRAOEAMLAIOLEKATKOELLTYIN	119
Db	87	LFSGHSAOGASTITQOULARNEFLS---PERTLMKRIEVEFLAIRIEDLTKDELTLEYLN	143
OY	120	KVYMSNGNOMQTAANOYVYKCDLNNLSLPTOLLAGMOPANOYDPSHEAODRRNIV	179
Db	144	KIYIGTAIYAGAAAOYTFKTYVDOLTLNEAYIATGPKPSTFNPLYSMDRAVARNIV	203
OY	180	LSEKKNQGYISADQYKAVNPTITDGLQSLKSSANPAYMDNLYKEVINQOE---ETGY	236
Db	204	LSRMLDEGYITQOQFDQTFREAIANAHAEIIFSAPLYISEMVRQEMYNRKYGSAVEDGY	263
OY	237	NLLITGMDVYTNVDQEO-----KH-----LMDIYN-----TDEYVA	268
Db	264	RIYTT---ITRKVOAAQOAVRNNVLDYDMRHGYRGFANLWVGESAMDNKITTDLKA	320
OY	269	YP-----DDEL	274
Db	321	LPYTGPLLPAAVTISANPQOATPMLADSGTVALSMEGVRMARPYRSDQOGPTPRKVTDL	380
OY	275	Q-----VASTIVDVS---NGKYIAOLGA---RHQSSNVSEGINQAV	309
Db	381	QTOGOIIVROYGDAMWLAQVPEVNSALVSINPOGAVMALVGGFEDFNQSK-----FNRA	435
OY	310	ETNDRMGSTMKPIYDPALEYGYDSTAIYHDEPYNYBGTNTPYVNMORG-----	365
Db	436	QALROVSSNIRKPEL-YTAAANDKCL--TLASMLND-----VPSRMDASAGSDMPK	483
OY	362	---YEGNITLOYALQOOSRNPVAVETLNVGLNRAKTFPLNGLIDPYSIHSNAISSNT	416
Db	484	NSPQOYAGPIRLRQGLGOSKNVWVVRAMRAGVDYAEYLOBRGCFPAQNIVHTESLA--	540
OY	417	TESKKTGASS---EKMAAAYAAFANGCYTKKPYIHKYVFSOG-----	457
Db	541	---LGSASFYPMOVARGYAAVANGSLVDPWFISKILNDGGVILEAKPKVACECD	594
OY	458	---SEKERSNV-----GTRAMKEETAYM--WTDMNKTYLVLYGTGRNAV-----	495
Db	595	IPVITGYDTQSNVLENNDDVEDVAISRQOQVNSVPMQLBEQANDALVAKTAQGEYAPHV	654
OY	496	-IAMLPOA-----GTTGTSNVTDEIEENHIKTSQF	524
Db	655	TPJAFLLTKSALNTNIFGEPGMOGTGWRBAGDRLQDRIDGKGTGTNN-----	699
OY	525	VAPDELPAQYTRKYSMAVWNTGYSNRLPLPLGVNGLTGYVAKYYSMMYTLSEGSNPEDMN	584
Db	700	SKSDAMVSSGIGPGVYISWVLGFDHRNL---GHTTASGAIKDOIISYEGGA-----	748
OY	585	EGLYRNGEYFKNGARSTWS-----PAPQOP--PSTESSSSSSDSSTQSSSTTPST	635
Db	749	---KSAOPAMDAIMKAVLEGVPEQPLTFPPPEIVTIVNIDRSTGO-----	788
OY	636	NNSTTNPNNNTOO---SNTTPDOO	657
Db	789	---LANGNSREYFIEGTQPTOO	809

TEST	RESULT	7
NO. 00 401		

```

US-08-481-435-10
; Sequence 10, Application US/08481435
Patent No. 6027906

GENERAL INFORMATION:
APPLICANT: Balganes, Tanjore S
APPLICANT: Town, Christine
TITLE OF INVENTION: No. 6027906el Polypeptides
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Steiner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 850 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
IMMEDIATE SOURCE:
CLONE: PARC0571 PPP 1A Q00A
US-08-481-435-10

```

Query Match	Score	DB 3;	Length
15.98;	555;		850;
Best Local Similarity	23.18;	Pred NO. 1	4e-33.

Matches 200; Conservative 129; Mismatches 246; Indels 290; Gaps 30;

```

QY 1 KIYNNKNOIADIGSERVANAQANDPITDLYKVAIVSIEDHRFPDHGIDITRL-GAFIR 59
Db 49 QIYSADGELIAOYGEKRRIVYTLDQIPBEWKAFLIEDSRFEHNGDVGITFRASVA 108
QY 60 NLOSNSILOGSGSTLTQOLIKITLYFESTSRDQITSRKAQEMALIALOLEKATKEILTYVIN 119
Db 109 LFSCHASQOGASTITPAALARNFPLS---PEKTLMRKIKEYVALIRIQELTKDEILLEYLN 165
QY 120 KYIWSNGNIGKQIAPNONTYKDKLNNISLPOLALLAGHPAOPNOYDPSHPEEAQODRRNLY 179
Db 166 KIYLYGRAYGGAQAQYFEKGTVDOLTLNEMAIYAGLPPKAPSTFENPLYSDMRVAVARRNV 225
QY 180 LSEKKNQOYISABOEKAVNTPITDGIQSLKASANYPAYMDNLTKEVINVEE---ETGY 236
Db 226 LSRMLADGYITOOOFDQITREALINANTHAEILPASFAYLISEWYROEKYNNYGGSAVEDGI 285

```



```

QY 237 NLLTGMVYTYNDQEAQ-----KH-----LMDIYN-----TDEYVA 268
DB 286 RIYTT---ITRRVQAAQAAQVARNVLDYDMRHGRCPANVLKVGSAADNNKITDTLKA 342
QY 269 YP-----DDLL 274
DB 343 LPTYPGLPLPAVTSANPQOATMLADGSTVALSMEGVWRARPYRSDTPQGPPIPRKVTDL 402
QY 275 Q-----VASTIVDS--NGKVIOLGA--RHSSNVSFIGNAV 309
DB 403 QTGQOIWROVQDAMWLAQVPEVNSALVSIINQNGAVMALVGGFENQSK-----FNRAI 457
QY 310 ETRNDGSTMKPIDTPAPALEXGVYDSTATIVHDEPNYNGTNPYNDRG----- 361
DB 458 QALROVGSNIKPEL-YTAMDKGL--TLASMLND-----VPISRMADASGSDWQPK 505
QY 362 -----YFCNITLOVALQOOSRNPVAVETLNKVLNRAKTFNLGLIDPYSIHSNAISSMT 416
DB 506 NSPPQYAGPIRLROGLGOSKNVVMRAMRGVDAEYLQRFPGFPAONIVHTESLA--- 562
QY 417 TESDKKYGAS--EKMAAFAAFANGCTYKKPMYIHKYVFDG----- 457
DB 563 -----LGSASFTPMQVAVMANGVFLVDPWFISKIENDQGVIFPAKPKVACPECD 616
QY 458 -----SEKEFSNV-----GTRAMKETTAYM-MTDMKTVLTYGTRNAY----- 495
DB 617 IYVIGDPOKSNVLENNDEVDVAISREOONVSVPMQLEQANOLVAKGADEYAPHVIN 676
QY 496 --LAMPQA-----GKTGTSNYTDEIENHKTSGF 524
DB 677 TPLAELIKSALNTNIFGEPMOGTGRAGRDILQRDIGKGTGTTN----- 721
QY 525 VAPDELFAGYTKRYKGMVATGYSNRLTPLYGNGLYAAVYVSMKTYLSEGSNPEDMNIP 584
DB 722 SKDWMFSQYGVVTSVWIGFDHRRNL--GHTTASGAILDQISGYEGGA----- 770
QY 585 EGLYNGEVEFNKRGARSTWNS-----PAPQOP--PSTESSSSSDSTSSSTTPTST 635
DB 771 -----KSAQOPAMDAYMAKVLGCVPEQPLPPPGIYTVNIDRSTGQ----- 810
QY 636 NNSTTTPNNNTQQ---SNTTPDQ 657
DB 811 -----LANGNSREYFIECTOPTQQ 831

```

RESULT 8

US-08-481-435-4

Sequence 4, Application US/08481435

Patent No. 6027906

GENERAL INFORMATION:

APPLICANT: Balganes, Tanjore S

APPLICANT: Town: Christine

TITLE OF INVENTION: No. 6027906el Polypeptides

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESS: White & Case

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2787

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/481,435

FILING DATE: 10-JUL-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: IN 580/MAS/94

```

; FILING DATE: 01-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9404072-2
; FILING DATE: 24-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 823 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-481-435-4

```

Query Match

15.1%; Score 526; DB 3; Length 823;

Best Local Similarity 28.9%; Pred No. 2,1e-31;

Matches 203; Conservative 99; Mismatches 273; Indels 128; Gaps 30;

```

QY 9 LIALDGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRIGAFRLNLS-NSLQ 67
DB 183 MISSPGEQRLVPRGSGFDLLVDLTATEDRHRYEHDGSIYSIGRAVLANTAGRTVQ 242
QY 68 GGSITLTOOLIKLTYSTSTSDTISKRAOEWLAIOLEKATKQELTYINKYV---- 123
DB 243 GASTLTOOLVKNLFUS---SERSYWRKANEAYMALIMDARYSDRLLELYMEVYLGQSG 299
QY 124 SNGNMGTAQONYGKDLNLSLPOLALLAGPOAPNQYDPSHPEAODRNLVLSM 183
DB 300 DNEIRGFPLASLYRGRPYEEELSDQALLVGVKASITYNWRNKALERRNLVRL 359
QY 184 KNOGYISAEQYKAVNPTPTDGLQSLKA-SNPAYMDNYLKEVI---NOVEEETGYNL 238
DB 360 QOOQIIDLQELYDMLSARPL--GVQPRGVISPPAPFMQLVROELQAKLDKVKYDLSGVI 417
QY 239 LTTGMDVYTYNDQEA-----QKHLMDIYNTDEYVAYDDDLQVASTIVDS 284
DB 418 FTT---FDSVAQDAEKAKEAVGIPALKRKRISD-----LETAIVVDRE 459
QY 285 NGKVIOLGARHQSNNVF--GINOAVETNRDNGSTMKPTDYAPALEYGVYDSTATIVHD 343
DB 460 SGEVRAMVG-----GSEPPQAGYNRRAMQARRSIGSLAKPAT-YLTALSQKTYIRLNTWIND 514
QY 344 EPNYNGTNPYVY---NMDRGY--EGNITLOVALQOOSRNPVAVETLNKVLNRAKTFNLG 398
DB 515 APIALRQPNQGVWSPONDRRYSESGRVMLVDALTRSMNVPTYNLGMALGLPAVETWIK 574
QY 399 LGIDPYSIHSNAI---SSNTTESDKKYGASSEKMAAATAANGTYKKPMYIHKYVS 455
DB 575 LGVPKDOLHPVAMLIGALNLTPIE-----VAQAQOTIASGGRAPLSALRSVIAE 625
QY 456 DGS--EKEFSNVGTRAMKETTYAMTDMKTYLTVTGTR--NAVYLAWLQAGTSGTNT 511
DB 626 DGVLYQSFPQ--AERAVPQAAVILITMWOVYQSTGOLGAKKYINLHLAKTGTN--- 682
QY 512 DEIENHIKTSQVAPDELFAGYTRKYSMAVVTGYSNRLTPLYGNGLYAAKVRSMYTY 571
DB 683 ---NNV-----DTMFAGIDGSTVITVWGRDNN-QPTKLYGASGAMSIYQ---RY 725
QY 572 LSEGSNPEDMNT--PE-----GLYRNGEVEFNKRGAR--STWNSPAPQOPSTESSSSSD 622
DB 726 LA-NQTPPLNLVPPEDIADMGVDYDGNFVCSGGMITLPWMS-----DPSLCOQ 775
QY 623 SSTSOSSTPTPNNTTTPNNNTQQSNTTTPDQONQONPOPO 665
DB 776 SEMQO-----QPSGNPFDQSSQPOQDPOQ-QPQD 803

```



```

: RESULT 9
: US-08-481-435-7
: Sequence 7, Application US/08481435
: Patent No. 6027906
: GENERAL INFORMATION:
: APPLICANT: Balanesh, Tanjore S
: APPLICANT: Town, Christine
: TITLE OF INVENTION: No. 6027906el Polypeptides
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: White & Case
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States
: ZIP: 10036-2787
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/481,435
: FILING DATE: 10-JUL-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: IN 580/MAS/94
: FILING DATE: 01-JUL-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: SE 9404072-2
: FILING DATE: 24-NOV-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Sterner, Richard J.
: REGISTRATION NUMBER: 35,372
: REFERENCE/DOCKET NUMBER: 1103326-151
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 819-8783
: TELEFAX: (212) 354-8113
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 844 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: ORIGINAL SOURCE:
: ORGANISM: Escherichia coli
: IMMEDIATE SOURCE:
: CLONE: PARC0438 PBP 1B QQAA
:
: US-08-481-435-7

```

```

Dh 439 FTT-----FDSTVAQDAAKRAKVEGIPALKKQRLSD-----LETAIVYVDRF 480
QY 285 NGKVI AQLGARHQSNNVSF-GINQAVETNRDMSGTKPITDYAPALEYGVYDSTATIVHD 343
Dh 481 SGCEVRAMWG-----GSEPOFAGYNRAMQARRSIGSLAKPAT-VLTALSQPKIYRLNFWIAD 535
QY 344 EPPYWPFGNTVY----NMDRKY--FGNITLQYALQCSRNPVAVETLKVGLNRAKPTFLNG 398
Dh 536 APIALROPNGCVSPQNDDRRYSESSEGRMLVDALTRSMNVPYVLGNALDLPVAVETWIK 595
QY 399 LGIYPSPIHYSNAI-----SNSTEDSKRYGASSEKMAAAYAFANGGYRYKPYTHKVVS 455
Dh 596 LGVRKDQLHPVPAMLGLNLTPIE-----VAAQFOTIASGGRAPLSALRSYIAE 646
QY 456 DGS--EKEFSNVGTRANKETTYAAMTDMKTVLTYGTR--NAVLAWLPOAGKTGTSNYT 511
Dh 647 DGKVLQSGFPQ-ABRAVPAQAAAYLTLLWMTMOOVGRGTGROLGAKYRPNLHLGKGTGN-- 703
QY 512 DELEIHNHTKSQFAVAPDELFAGYTRKYSMAWVTGYSNRLLPYVNGITVAAKYRSMY 571
Dh 704 -----NNV-----DTWFAIDGISTVYTTWGRDNN-OPTKLYGASGAMSTYQ--RY 746
QY 572 LSEGSNPEDWNI--PE-----GLYRNGEYFKNCAR--STWNSPAPQPPSTESSSSD 622
Dh 747 LA-NOTPPPLNLYVPEDIDADMGVYIDGNFVCSGSMRLLPWTSS-----DPQSICQ 796
QY 623 SSTQSSSTPSTNNSTYTTNPNNNTQOASNTYTPDOQONQPPAQ 665
Dh 797 SEMQ-----QPSGNPPOSSQPPQPPQ-QPAQ 824

RESULT 10
US-08-481-435-8
Sequence 8, Application US/08481435
Patent No. 6027906
GENERAL INFORMATION:
APPLICANT: Balgatesh, Tanjore S
APPLICANT: Town, Christine
TITLE OF INVENTION: No. 6027906el Polypeptides
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 110326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 844 amino acids

```


TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Escherichia coli
 IMMEDIATE SOURCE:
 CLONE: PARC0468 PBP 1B QOOL
 US-08-481-435-8

Query Match 14.7%; Score 512; DB 3; Length 844;
 Best Local Similarity 28.6%; Pred. No. 2.5e-30;
 Matches 201; Conservative 99; Mismatches 275; Indels 128; Gaps 30;

```

QY 9 L I A D L G S E R R V N A Q A N D I P T D L V K A I V S T E D H R F E D H R G I D T R I L G A F L R N L Q S N S L Q 67
DB 204 M I S S P N G E Q R L F V P R S G F P D L V D T L L A T E D R H F E H D I S I S I G R A V I A N L T A G R T V Q 263
QY 68 G G S T L T Q O L I K L T Y F S T S D Q T I S R K A O E A M L A I O L E K A T R K O E I L T Y I N K V Y M - - - - 123
DB 264 G A S T L F L L V K N L F L S - - - S E R S Y W R K A N E A Y A L L I M D A R Y S K D R I L E L Y M N E Y I L Q S G 320
QY 124 S N C Y G M Q T A A O N Y Y K G K D L N N L S L P O L A L L A G P Q A P N O Y D P Y S H P E A A O D R R N L V S E M 183
DB 321 D M E I R G F P L A S I Y F G R P E E L S I D O A L L V G W K G A S I Y N P R N P K L A L E R N L V L R L L 380
QY 184 K N G Y I S A B O Y E K A V N T P I T D G L Q S L K S A - S N Y P A Y M D N L K E V I - - - N O V E E T G Y N L 238
DB 381 Q O O O I I D O E L Y D M L S A R P L - G V Q P R G V I S P O P A F M O L V R O L Q A K L G D K V D L S G V K I 438
QY 239 L T T G M D V Y T N V D O E A - - - - - O K H L M D I Y N T D E V A Y P D D E L O V A S T I V D S 284
DB 439 F T T - - - F D S V A O D A E K A A V E G I P A L K K R K L S D - - - - - L E T A I V V D R F 480
QY 285 N E K V I A Q L G A R H O S S V N S F - G I N O A V E T N R D M G S T M K P I T D Y A P A L E Y G Y D S T A T I V H D 343
DB 481 S G E V R A M V G - - - - G S E P Q F G Y N R A M Q A R R S I G S I A K P A T - Y L T A L S O P K I Y R L N T W I A D 535
QY 344 E P Y N V P G I T N T P V Y - - - N M D R G Y - - F G N I T L O Y A L Q O S R N V P A V E T L N K Y G L N R A K T F L N G 398
DB 536 A P I A L R Q P R G O W S P O N D R R Y S E S G R V M L V D A L T R S M V P Y N L G M A L G I P A V T E T W I K 595
QY 399 L G I D P S I H Y S N A I - - - S S N T E S D K K Y G A S S E K M A A Y A A P A N G S T Y Y K P M Y I H K V F S 455
DB 596 L G V P K D Q L P H P A M L L G A L N L P I E - - - - - V A O F Q T I A S G N R A P L S A L R S V I E 646
QY 456 D S S - E K E S N N G T R A M K E T T A Y A M T D M K Y I L T Y T G R - N A Y L A M L F O A Q T G I S N T 511
DB 647 D G K V L Y O S F P Q - A E R A V P A Q A A L L T L M T M O Q V V O R G T G R O L G A K Y P N L H L A G K T G T T N - 703
QY 512 D E I E N H I T S O F V A P D E L F A G Y T R K Y S M A V N T G Y S N R L T P L A V G N L J V A A K Y R S M M T Y 571
DB 704 - - - - - N N V - - - - - D T M F A G I D S T V I T W Y G R D N N - Q P R K L Y A S A S A M S I Y Q - - R Y 746
QY 572 L E E G S P E E M N T - - P E - - - - - G L Y N G E F V F K N G A R - - S T W S P A P O Q P P S T E S S S S S D 622
DB 747 I A - N O T P P L N I V P P E D I A D M G V D Y D G N F V C S G M R I L P V M T S - - - - - D P Q S L C O Q 796
QY 623 S T S O S S S T P T S N S T T T N P N N N T Q O S N T P D Q O N Q O P A Q 665
DB 797 S E M O Q - - - - - - - - - - - Q P S G N P P D O S S P O Q O P Q - Q P A Q 824
  
```

RESULT 11
 US-08-481-435-9
 ; Sequence 9, Application US/08481435
 ; Patent No. 6027906
 ; GENERAL INFORMATION:
 ; APPLICANT: Balganes, Tanjore S
 ; APPLICANT: Town, Christine
 ; TITLE OF INVENTION: No. 6027906el Polypeptides
 ; NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:
 ADDRESSEE: White & Case
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States
 ZIP: 10036-2787
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/481,435
 FILING DATE: 10-JUL-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: IN 580/MAS/94
 FILING DATE: 01-JUL-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9404072-2
 FILING DATE: 24-NOV-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Steiner, Richard J.
 REGISTRATION NUMBER: 35,372
 REFERENCE/DOCKET NUMBER: 1103326-151
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 819-8783
 TELEFAX: (212) 354-8113
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 836 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Escherichia coli
 IMMEDIATE SOURCE:
 CLONE: PARC0469 PBP 1B del 8
 US-08-481-435-9

Query Match 14.0%; Score 489; DB 3; Length 836;
 Best Local Similarity 28.1%; Pred. No. 1.3e-28;
 Matches 197; Conservative 97; Mismatches 274; Indels 134; Gaps 30;

```

QY 9 L I A D L G S E R R V N A Q A N D I P T D L V K A I V S T E D H R F E D H R G I D T R I L G A F L R N L Q S N S L Q 68
DB 204 M I S S P N G E Q R L F V P R S G F P D L V D T L L A T E D R H F E H D I S I S I G R A V I A N L T A G R - - - 260
QY 69 G S T L T Q O L I K L T Y F S T S D Q T I S R K A O E A M L A I O L E K A T R K O E I L T Y I N K V Y M - - - - S 124
DB 261 - - - - - I Y Q L V K N L F L S - - - S E R S Y W R K A N E A Y A L L I M D A R Y S K D R I L E L Y M N E Y I L Q S G D 313
QY 125 N C Y G M Q T A A O N Y Y K G K D L N N L S L P O L A L L A G P Q A P N O Y D P Y S H P E A A O D R R N L V S E M K 184
DB 314 N E I R G F P L A S I Y F G R P E E L S I D O A L L V G W K G A S I Y N P R N P K L A L E R N L V L R L L 373
QY 185 N G Y I S A B O Y E K A V N T P I T D G L Q S L K S A - S N Y P A Y M D N L K E V I - - - N O V E E T G Y N L 239
DB 374 Q O O O I I D O E L Y D M L S A R P L - G V Q P R G V I S P O P A F M O L V R O L Q A K L G D K V D L S G V K I F 431
QY 240 T T G M D V Y T N V D O E A - - - - - O K H L M D I Y N T D E V A Y P D D E L O V A S T I V D S N 285
DB 432 T T - - - F D S V A O D A E K A A V E G I P A L K K R K L S D - - - - - L E T A I V V D R F S 473
QY 286 G V I A Q L G A R H O S S V N S F - G I N O A V E T N R D M G S T M K P I T D Y A P A L E Y G Y D S T A T I V H D E 344
DB 474 G E Y R A M V G - - - - G S E P Q F G Y N R A M Q A R R S I G S I A K P A T - Y L T A L S O P K I Y R L N T W I A D A 528
QY 345 P Y N P E T N T P V Y - - - N M D R G Y - - F G N I T L O Y A L Q O S R N V P A V E T L N K Y G L N R A K T F L N G 399
  
```



```

Db 529 PIALQPGQVWSPQNDRRRSESGRWLVDALTRSNVPTVNLGALCPAVETWIKL 588
Qy 400 GIDYPSIHYNAI--SSNTTESDKKYGASSEKMAAAYAFANGSTYKKPMYIHKVPSD 456
Db 589 GVPKQQLPVPAMLIGALNLTPIE-----VAQAFOTIASGGRAPLSALRVED 639
Qy 457 GS--EKESNNGTRAMKETTYAMTDMKKIVLYTGTGR--NAYLAWLPQAGTGTSTNTD 512
Db 640 GKVLQSPFO--AERAVPQAAAYLLTMTMQOVORCTGRQLGAKYFNHLAKGTGTN--- 695
Qy 513 EEIENHITSGFVAPDELFCATRYKRYSAVMYGYSNRLPLVNGNLTYAAKYRSMYTL 572
Db 656 -----NNV-----DIWFAGIDGSTVTITWVGRDN--QPTKLYGASGAMSITQ---RYL 739
Qy 573 SEGSPEDMNI--PE-----GLYNGGEVFFKNGAR--STWNSPAPQOPSTESSSSSDS 623
Db 740 A-NQTPPLNLVPPEDIDMGVDYDGNVFCGGMRLPLVWIS-----DPSLCOQS 789
Qy 624 STSQSSSTPSTNNSTTNPNNNTQOQNTPTPDQONQNPQPAQ 665
Db 790 EMOQ-----QPSGNPFDOSSQPOQPOQ-OPAQ 816

```

```

RESULT 12
US-08-245-511-24
: Sequence 24, Application US/08245511
: Patent No. 5928900
: GENERAL INFORMATION:
: APPLICANT: Masure, H Robert
: APPLICANT: Pearce, Barbara J
: APPLICANT: Tuomanen, Elaine
: TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
: NUMBER OF SEQUENCES: 58
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Klauber & Jackson
: STREET: 411 Hackensack Avenue
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/245,511
: FILING DATE: 18-MAY-1994
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/116,541
: FILING DATE: 01-SEP-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE/DOCKET NUMBER: 600-1-069 CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201 487-5800
: TELEFAX: 201 343-1684
: TELEX: 133521
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 77 amino acids
: TYPE: amino acid
: TOPOLOGY: unknown
: MOLECULE TYPE: peptide
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: Streptococcus pneumoniae
: STRAIN: R6

```

```

: IMMEDIATE SOURCE:
: CLONE: SPRU42
: US-08-245-511-24

```

```

Query Match 11.7%; Score 406; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 6,2e-24;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 240 TTGMDVYTNVDOEAKHMDIYNTEDEYVAYPPDELOVASTIYDVSNGKVIQOLGARHOSS 299
Db 1 TTGMDVYTNVDOEAKHMDIYNTEDEYVAYPPDELOVASTIYDVSNGKVIQOLGARHOSS 60
Qy 300 NVSEGINQAVETNRDNG 316
Db 61 NVSEGINQAVETNRDNG 77

```

```

RESULT 13
US-08-600-993A-24
: Sequence 24, Application US/08600993A
: Patent No. 5981229
: GENERAL INFORMATION:
: APPLICANT: Masure, H Robert
: APPLICANT: Pearce, Barbara J
: APPLICANT: Tuomanen, Elaine
: TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
: NUMBER OF SEQUENCES: 59
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Klauber & Jackson
: STREET: 411 Hackensack Avenue
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/600,993A
: FILING DATE: 1-MAR-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/245,511
: FILING DATE: 18-MAY-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/116,541
: FILING DATE: 01-SEP-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE/DOCKET NUMBER: 600-1-069 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201 487-5800
: TELEFAX: 201 343-1684
: TELEX: 133521
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 77 amino acids
: TYPE: amino acid
: TOPOLOGY: unknown
: MOLECULE TYPE: peptide
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: Streptococcus pneumoniae
: STRAIN: R6
: IMMEDIATE SOURCE:

```


CLONE: SPRU42
US-08-600-993A-24

Query Match 11.7%; Score 406; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 6.2e-24;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 240 TTGMDVYTNVDEDAQKHLMDIYNTEDEYVAYPPDELOVASTIVDSNGKYIAQLGARHSS 299
Db 1 TTGMDVYTNVDEDAQKHLMDIYNTEDEYVAYPPDELOVASTIVDSNGKYIAQLGARHSS 60

Qy 300 NVSEGINAVETNRDMG 316
Db 61 NVSEGINAVETNRDMG 77

RESULT 14
US-08-481-435-12
; Sequence 12, Application US/08481435
; Patent No. 6027906
; GENERAL INFORMATION:
; APPLICANT: Balganes, Tanjore S
; APPLICANT: Town, Christine
; TITLE OF INVENTION: No. 6027906el Polypeptides
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,435
; FILING DATE: 10-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IN 580/MAS/94
; FILING DATE: 01-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9404072-2
; FILING DATE: 24-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Steiner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; IMMEDIATE SOURCE:
; CLONE: PARC 0593 truncated soluble PBP 1B
US-08-481-435-12

Query Match 10.6%; Score 370.5; DB 3; Length 532;
Best Local Similarity 30.7%; Pred. No. 6.2e-20;
Matches 115; Conservative 68; Mismatches 138; Indels 53; Gaps 12;

Qy 9 LIADLGSERRVNAQANDIPTLVKALVSTIEDHREFDHRCIDPTIRILGAFLRNLSQ-NSLQ 67
Db 183 MISPPNGEQLRLEVPYRSGFDPLLVDLTATEDRHHFYEHDCISLSTIGRAVLNLTAGRTVQ 242

Qy 68 GGSSTLTLQOLIKLTYPSTISDOTISKRAQEAHLALOLEKATKQELTLYINKVYM----- 123
Db 243 GASTLTLQOLVKKLFLS-----SERSYWRKANEAYMALIMDARYSKDRLLLEYMNEVYLGSG 299

Qy 124 SNGNYGMDTAQNNYKGDINNLSLPOLALLAGMPQAPNODPYSHPEAQDNRNLVSEM 183
Db 300 DNEIRGFPLASLYRFRPVEELSLDQALLVGMKASLYNWRNRKALDERNLVRL 359

Qy 184 KNOGYTSAEQYKAVNTPITDGLSLKSA-SNPAYMDNYLKEVI---NOVEETGYWL 238
Db 360 OOOQILIDQELYLMLSARPL-GVQPRGVISPPAFMQLVRQELQAKLGDKVKDLSGVKXI 417

Qy 239 LTGMDVYTNVDEA-----QKHLMDIYNTEDEYVAYPPDELOVASTIVDS 284
Db 418 FTT---FDSVADAAEKAAVEGIPALKKORKLSD-----LETAIVVDHR 459

Qy 285 NGKYIAQLGARHSSNVSE-GINAVETNRDMGSTKPKPTDYAPALEYGVDSSTATIVHD 343
Db 460 SEVRRAMVG-----GSEPPQAGYNRAMQARSTIGSLAKPAT-ITLALSQPKYIRLNTWIND 514

Qy 344 EPNYNPGETNTPVYN 357
Db 515 APIALRQPNQVWS 528

RESULT 15
US-08-481-435-11
; Sequence 11, Application US/08481435
; Patent No. 6027906
; GENERAL INFORMATION:
; APPLICANT: Balganes, Tanjore S
; APPLICANT: Town, Christine
; TITLE OF INVENTION: No. 6027906el Polypeptides
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,435
; FILING DATE: 10-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IN 580/MAS/94
; FILING DATE: 01-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9404072-2
; FILING DATE: 24-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Steiner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:40:34 ; Search time 115.35 Seconds

(Without alignments)
2032.241 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484

Sequence: 1 KIYDNKNOIADLGSRVYN.....TQSSNTTPPOQNONPQAPQ 666

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCNUS_COMB.pep.*

2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*

3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*

4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*

5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*

6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*

7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*

8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*

9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*

10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*

11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*

12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*

13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*

14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*

15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*

16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*

17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*

18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*

19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*

20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*

21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*

22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*

23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*

24: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*

25: /cgn2_6/ptodata/2/paa/US101_COMB.pep.*

26: /cgn2_6/ptodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3484	100.0	666	19	US-09-536-784-2
2	3484	100.0	666	21	US-09-765-271-2
3	3484	100.0	666	21	US-09-765-272-2
4	3483	100.0	719	19	US-09-583-110-3973
5	3483	100.0	721	15	US-09-107-433-3705
6	3466	99.5	719	1	PCT-US02-03987-13423
7	3466	99.5	719	22	US-09-815-242-13423

8	3466	99.5	719	24	US-10-072-851-13423	Sequence 13423, A
9	1862.5	53.5	823	15	US-09-107-532-5667	Sequence 5567, Ap
10	1862.5	53.5	823	15	US-09-107-532A-5667	Sequence 5667, Ap
11	1751	50.3	778	1	PCT-US02-03987-10728	Sequence 10728, A
12	1751	50.3	778	22	US-09-815-242-10728	Sequence 10728, A
13	1751	50.3	778	24	US-10-072-851-10728	Sequence 10728, A
14	1751	50.3	789	15	US-09-134-000-4939	Sequence 4939, Ap
15	1621	46.5	320	5	US-08-116-541-4	Sequence 4, Appl1
16	1473.5	42.3	771	20	US-09-634-238-229	Sequence 229, App
17	1473.5	42.3	771	20	US-09-634-238-262	Sequence 262, App
18	1361	39.1	462	24	US-10-091-007-92	Sequence 92, Appl
19	817.5	23.5	727	1	PCT-US02-03987-5782	Sequence 5782, Ap
20	817.5	23.5	727	1	PCT-US02-03987-12777	Sequence 12777, A
21	817.5	23.5	727	22	US-09-815-242-5782	Sequence 5782, Ap
22	817.5	23.5	727	22	US-09-815-242-12777	Sequence 12777, A
23	817.5	23.5	727	24	US-10-072-851-5782	Sequence 5782, Ap
24	817.5	23.5	727	20	US-10-072-851-12777	Sequence 12777, A
25	815.5	23.4	730	20	US-09-611-529-4512	Sequence 4512, Ap
26	811.5	23.3	716	26	US-60-253-625-2437	Sequence 2437, Ap
27	811.5	23.3	716	26	US-60-257-931-2821	Sequence 2821, Ap
28	811.5	23.3	716	26	US-60-269-308-3840	Sequence 3840, Ap
29	809.5	23.2	746	15	US-09-134-001C-3214	Sequence 3214, Ap
30	661.5	19.0	675	20	US-09-634-238-228	Sequence 228, App
31	655	18.8	698	21	US-09-739-449-11825	Sequence 11825, A
32	655	18.8	698	22	US-09-803-110-11825	Sequence 11825, A
33	648.5	18.6	814	17	US-09-328-352-4373	Sequence 4373, Ap
34	641	18.4	834	16	US-09-252-991A-32671	Sequence 32671, A
35	636	18.3	730	15	US-09-107-433-4707	Sequence 4707, Ap
36	636	18.3	731	1	PCT-US97-19070-2	Sequence 2, Appl1
37	636	18.3	731	19	US-09-583-110-4720	Sequence 4720, Ap
38	632.5	18.2	855	17	US-09-328-352-6216	Sequence 6216, Ap
39	625	17.9	660	1	PCT-US97-05223-845	Sequence 845, App
40	625	17.9	660	1	PCT-US97-05223-876	Sequence 876, App
41	625	17.9	660	11	US-08-761-318-845	Sequence 845, App
42	625	17.9	660	11	US-08-761-318-876	Sequence 876, App
43	625	17.9	660	12	US-08-824-132-845	Sequence 845, App
44	625	17.9	660	12	US-08-824-132-876	Sequence 876, App
45	625	17.9	660	13	US-08-993-002A-5603	Sequence 5603, Ap

ALIGNMENTS

RESULT 1

US-09-536-784-2

Sequence 2, Application US/09536784

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/536,784

FILING DATE: 30-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/961,083

FILING DATE: OCT-30-1997

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB34063


```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-536-784-2

Query Match      100.0%; Score 3484; DB 19; Length 666;
Best Local Similarity 100.0%; Pred. No. 3e-289;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNQIADLGSERRNAQAANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLELN 60
DB 1 KIYDNKNQIADLGSERRNAQAANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLELN 60
QY 61 LOSNSLQGGSTLTLOOLIKLTFSTSDQTSRKQAEAMLAIOLEOKATKOELITYYINK 120
DB 61 LOSNSLQGGSTLTLOOLIKLTFSTSDQTSRKQAEAMLAIOLEOKATKOELITYYINK 120
QY 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAADRNLVL 180
DB 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAADRNLVL 180
QY 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAADRNLVL 180
DB 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAADRNLVL 180
QY 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLLT 240
DB 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLLT 240
QY 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLLT 240
DB 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLLT 240
QY 241 TGMADVYTNVDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDVSNKVIQAOLGARQSSN 300
DB 241 TGMADVYTNVDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDVSNKVIQAOLGARQSSN 300
QY 301 VSEFGINQAVETNRDMGSTMKPITDYAPALEYGYDSTATIVHDEPNYPGTNTPVYNNMDR 360
DB 301 VSEFGINQAVETNRDMGSTMKPITDYAPALEYGYDSTATIVHDEPNYPGTNTPVYNNMDR 360
QY 361 GYFGNITLQYALQOQSRNVPAVELLNKVGILNRAKTFPLNGCIDYPSIHYSNALSSNTTESD 420
DB 361 GYFGNITLQYALQOQSRNVPAVELLNKVGILNRAKTFPLNGCIDYPSIHYSNALSSNTTESD 420
QY 421 KYYGASSEKMAAAYAFANGGTYIKPMYIHKYVFSGSEKESFNSVGT RAMKETTAIAMD 480
DB 421 KYYGASSEKMAAAYAFANGGTYIKPMYIHKYVFSGSEKESFNSVGT RAMKETTAIAMD 480
QY 481 MMKTYLTYGTGRNAYLAWLPQAGKTGTSNYTDEEIEENHIKTSQFVAPDELFAGYTRKYSM 540
DB 481 MMKTYLTYGTGRNAYLAWLPQAGKTGTSNYTDEEIEENHIKTSQFVAPDELFAGYTRKYSM 540
QY 541 AWTGCGSNLTPLVNGCLTVAAYKVRSMATYLSGGSNPEDMNPBGLYNGEVEFGNGAR 600
DB 541 AWTGCGSNLTPLVNGCLTVAAYKVRSMATYLSGGSNPEDMNPBGLYNGEVEFGNGAR 600
QY 601 STWNSBAPQPPSTESSSSSDSSTPSTNNSTTTNNNTTQOQNTTPOOQON 660
DB 601 STWNSBAPQPPSTESSSSSDSSTPSTNNSTTTNNNTTQOQNTTPOOQON 660
QY 661 PQAPQ 666
DB 661 PQAPQ 666
```

```
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,271
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/536,784
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/961,083
FILING DATE: OCT-30-1997
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-765-271-2

Query Match      100.0%; Score 3484; DB 21; Length 666;
Best Local Similarity 100.0%; Pred. No. 3e-289;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNQIADLGSERRNAQAANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLELN 60
DB 1 KIYDNKNQIADLGSERRNAQAANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLELN 60
QY 61 LOSNSLQGGSTLTLOOLIKLTFSTSDQTSRKQAEAMLAIOLEOKATKOELITYYINK 120
DB 61 LOSNSLQGGSTLTLOOLIKLTFSTSDQTSRKQAEAMLAIOLEOKATKOELITYYINK 120
QY 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAADRNLVL 180
DB 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAADRNLVL 180
QY 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLLT 240
DB 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLLT 240
QY 241 TGMADVYTNVDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDVSNKVIQAOLGARQSSN 300
DB 241 TGMADVYTNVDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDVSNKVIQAOLGARQSSN 300
QY 301 VSEFGINQAVETNRDMGSTMKPITDYAPALEYGYDSTATIVHDEPNYPGTNTPVYNNMDR 360
DB 301 VSEFGINQAVETNRDMGSTMKPITDYAPALEYGYDSTATIVHDEPNYPGTNTPVYNNMDR 360
QY 361 GYFGNITLQYALQOQSRNVPAVELLNKVGILNRAKTFPLNGCIDYPSIHYSNALSSNTTESD 420
DB 361 GYFGNITLQYALQOQSRNVPAVELLNKVGILNRAKTFPLNGCIDYPSIHYSNALSSNTTESD 420
QY 421 KYYGASSEKMAAAYAFANGGTYIKPMYIHKYVFSGSEKESFNSVGT RAMKETTAIAMD 480
```



```
|||||
Db 421 KKGASSEKMAAAVAAPANGTYYKPMYIKHVPSDSEKEEFSNVGRAMKETTAYAMTD 480
QY 481 MKTIVLTGGRNAYLAMLPOAGKTGTSNTDEIEHNIKTSQVADDELFACTRRYSM 540
Db 481 MKTIVLTGGRNAYLAMLPOAGKTGTSNTDEIEHNIKTSQVADDELFACTRRYSM 540
QY 541 AVMTGYNRLPLVGNGLTVAAKYRRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
|||||
Db 541 AVMTGYNRLPLVGNGLTVAAKYRRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
QY 601 STWNSPAPQOPPTSESSSSSDSSTQSSSTPSTNNSTTNPNNNTQOQNTTDDQONON 660
|||||
Db 601 STWNSPAPQOPPTSESSSSSDSSTQSSSTPSTNNSTTNPNNNTQOQNTTDDQONON 660
QY 661 POPAOP 666
|||||
Db 661 POPAOP 666

RESULT 3
; Sequence 2, Application US/09765272
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-765-272-2.

Query Match 100.0%; Score 3484; DB 21; Length 666;
Best Local Similarity 100.0%; Pred. No. 36-289;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
|||||
Db 61 LOSNSLGGSTLPLQOLIKLTYFTSTSDQTSRKAQAWLAIOLEKATKOEILTYINK 120
QY 121 VYNSNGNYGMOATAQNTYGGKDLNLSLPOLALLAGMOAPQNPQYPSHPPEAQQRRLLYL 180
Db 121 VYNSNGNYGMOATAQNTYGGKDLNLSLPOLALLAGMOAPQNPQYPSHPPEAQQRRLLYL 180
QY 181 SEMNNGYISAEOYEKAVNMPITFDGLSLKSASNPAYMDNYLKEVINQVEEFTGYMLLT 240
|||||
Db 181 SEMNNGYISAEOYEKAVNMPITFDGLSLKSASNPAYMDNYLKEVINQVEEFTGYMLLT 240
QY 241 TGMQVYTNVDQEAQKHLMDIYNMDEVYAYPDDELQVASTIVDVSNGKVIQLGARHSSN 300
|||||
Db 241 TGMQVYTNVDQEAQKHLMDIYNMDEVYAYPDDELQVASTIVDVSNGKVIQLGARHSSN 300
QY 301 VSFGINQAVETNRDWSGTMKRPITDYAPALEGYDSTATTIVHDEPYNYPGNTFTVYWMDR 360
|||||
Db 301 VSFGINQAVETNRDWSGTMKRPITDYAPALEGYDSTATTIVHDEPYNYPGNTFTVYWMDR 360
QY 361 GYEGNITLOYALQOSRNVPAVETLNKVLNRAKTFELNGLDIYPSIHSNAISSNTTESD 420
|||||
Db 361 GYEGNITLOYALQOSRNVPAVETLNKVLNRAKTFELNGLDIYPSIHSNAISSNTTESD 420
QY 421 KKGASSEKMAAAVAAPANGTYYKPMYIKHVPSDSEKEEFSNVGRAMKETTAYAMTD 480
|||||
Db 421 KKGASSEKMAAAVAAPANGTYYKPMYIKHVPSDSEKEEFSNVGRAMKETTAYAMTD 480
QY 481 MKTIVLTGGRNAYLAMLPOAGKTGTSNTDEIEHNIKTSQVADDELFACTRRYSM 540
|||||
Db 481 MKTIVLTGGRNAYLAMLPOAGKTGTSNTDEIEHNIKTSQVADDELFACTRRYSM 540
QY 541 AVMTGYNRLPLVGNGLTVAAKYRRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
|||||
Db 541 AVMTGYNRLPLVGNGLTVAAKYRRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
QY 601 STWNSPAPQOPPTSESSSSSDSSTQSSSTPSTNNSTTNPNNNTQOQNTTDDQONON 660
|||||
Db 601 STWNSPAPQOPPTSESSSSSDSSTQSSSTPSTNNSTTNPNNNTQOQNTTDDQONON 660
QY 661 POPAOP 666
|||||
Db 661 POPAOP 666

RESULT 4
; US-09-583-110-3973
; Sequence 3973, Application US/09583110
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3973
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3973

Query Match 100.0%; Score 3483; DB 19; Length 719;
Best Local Similarity 99.8%; Pred. No. 4,2e-289;
Matches 665; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```


|||||
Db 54 KIYNNKNLIDLGSERRVNAQANDIPDLVKAIVSIEDHFRFDRGIDTIRILGAFRLN 113
QY 61 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSRKAQOAMLAIOLEOKATKOEILLYYINK 120
Db 114 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSRKAQOAMLAIOLEOKATKOEILLYYINK 173
QY 121 VYMSNGNGMGTAAOANYGKDLNNLSLPQALLAGMPAPNOYDPYSHPEAAQDRRLNVL 180
Db 174 VYMSNGNGMGTAAOANYGKDLNNLSLPQALLAGMPAPNOYDPYSHPEAAQDRRLNVL 233
QY 181 SEMKNQGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNYLKEVINQVEEETYNLLT 240
Db 234 SEMKNQGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNYLKEVINQVEEETYNLLT 293
QY 241 TGMDEVYTNVDOAKHMDIYNTDEYVAYPPDELOVASTIVDVSNKGKXIAQIGARHOSSN 300
Db 294 TGMDEVYTNVDOAKHMDIYNTDEYVAYPPDELOVASTIVDVSNKGKXIAQIGARHOSSN 353
QY 301 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGTNTPYNNDR 360
Db 354 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGTNTPYNNDR 413
QY 361 GYFGNITLQYALQOSRNPAVETLNKVGILNRKTFPLNGLGIDYPSIHYSNALSSNTTESD 420
Db 414 GYFGNITLQYALQOSRNPAVETLNKVGILNRKTFPLNGLGIDYPSIHYSNALSSNTTESD 473
QY 421 KKYGSSSKMAAAYAFANGGTYRPMYIHKYVFSGSEKESNNGTAMKETTAYMMTD 480
Db 474 KKYGSSSKMAAAYAFANGGTYRPMYIHKYVFSGSEKESNNGTAMKETTAYMMTD 533
QY 481 MKKTVLTGTRNAYLANLPOAGKTGTSNYTDEELENHIKTSQFVAPDELFAGYTRKYSM 540
Db 534 MKKTVLTGTRNAYLANLPOAGKTGTSNYTDEELENHIKTSQFVAPDELFAGYTRKYSM 593
QY 541 AVMGTYSNRLLPVLVNGLTVAAKYRSMWTYLSGSSNPEDMNIPEGILRNGEFVRKNGAR 600
Db 594 AVMGTYSNRLLPVLVNGLTVAAKYRSMWTYLSGSSNPEDMNIPEGILRNGEFVRKNGAR 653
QY 601 STWNSPAPQOPSTSSSSSSSDSSSTPSTNNSTTTNPNNNTOOSNTTPOOQON 660
Db 654 STWNSPAPQOPSTSSSSSSSDSSSTPSTNNSTTTNPNNNTOOSNTTPOOQON 713
QY 661 POPAP 666
Db 714 POPAP 719

RESULT 5
US-09-107-433-3705
Sequence 3705, Application US/09107433
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-stramm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS FOR DIAGN
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <unknown>
OPERATING SYSTEM: <unknown>
SOFTWARE: <unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Atinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3705:
SEQUENCE CHARACTERISTICS:
LENGTH: 721 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...721
SEQUENCE DESCRIPTION: SEQ ID NO: 3705:
US-09-107-433-3705
Query Match 100.0%; Score 3483; DB 15; Length 721;
Best Local Similarity 99.8%; Pred. No. 4,3e-289;
Matches 665; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KIYNNKNLIDLGSERRVNAQANDIPDLVKAIVSIEDHFRFDRGIDTIRILGAFRLN 60
Db 56 KIYNNKNLIDLGSERRVNAQANDIPDLVKAIVSIEDHFRFDRGIDTIRILGAFRLN 115
QY 61 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSRKAQOAMLAIOLEOKATKOEILLYYINK 120
Db 116 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSRKAQOAMLAIOLEOKATKOEILLYYINK 175
QY 121 VYMSNGNGMGTAAOANYGKDLNNLSLPQALLAGMPAPNOYDPYSHPEAAQDRRLNVL 180
Db 176 VYMSNGNGMGTAAOANYGKDLNNLSLPQALLAGMPAPNOYDPYSHPEAAQDRRLNVL 235
QY 181 SEMKNQGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNYLKEVINQVEEETYNLLT 240
Db 236 SEMKNQGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNYLKEVINQVEEETYNLLT 295
QY 241 TGMDEVYTNVDOAKHMDIYNTDEYVAYPPDELOVASTIVDVSNKGKXIAQIGARHOSSN 300
Db 296 TGMDEVYTNVDOAKHMDIYNTDEYVAYPPDELOVASTIVDVSNKGKXIAQIGARHOSSN 355
QY 301 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGTNTPYNNDR 360
Db 356 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGTNTPYNNDR 415
QY 361 GYFGNITLQYALQOSRNPAVETLNKVGILNRKTFPLNGLGIDYPSIHYSNALSSNTTESD 420
Db 416 GYFGNITLQYALQOSRNPAVETLNKVGILNRKTFPLNGLGIDYPSIHYSNALSSNTTESD 475
QY 421 KKYGSSSKMAAAYAFANGGTYRPMYIHKYVFSGSEKESNNGTAMKETTAYMMTD 480
Db 476 KKYGSSSKMAAAYAFANGGTYRPMYIHKYVFSGSEKESNNGTAMKETTAYMMTD 535
QY 481 MKKTVLTGTRNAYLANLPOAGKTGTSNYTDEELENHIKTSQFVAPDELFAGYTRKYSM 540
Db 536 MKKTVLTGTRNAYLANLPOAGKTGTSNYTDEELENHIKTSQFVAPDELFAGYTRKYSM 595
QY 541 AVMGTYSNRLLPVLVNGLTVAAKYRSMWTYLSGSSNPEDMNIPEGILRNGEFVRKNGAR 600
Db 596 AVMGTYSNRLLPVLVNGLTVAAKYRSMWTYLSGSSNPEDMNIPEGILRNGEFVRKNGAR 655
QY 601 STWNSPAPQOPSTSSSSSSSDSSSTPSTNNSTTTNPNNNTOOSNTTPOOQON 660

Db 656 STWNSPAPQPPSTRESSSSSDSTSOSSTPTTNNNTTTPNNNTTTPDOONON 715
Qy 661 POPAOP 666
Db 716 POPAOP 721

RESULT 6
PCT-US02-03987-13423

; Sequence 13423, Application PC/WUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits C
; FILE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; PRIOR FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
PCT-US02-03987-13423

Query Match 99.5%; Score 3466; DB 1; Length 719;
Best Local Similarity 99.4%; Pred. No. 1.2e-287;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KIYDKRKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLENN 60
Db 54 KIYDKRKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLENN 113
Qy 61 LOSNSLOGSSTLTQOLIKLTYFSTSTSDQTSRKAQEAAMLQLEOKATKQELITYYINK 120
Db 114 LOSNSLOGSALTQOLIKLTYFSTSTSDQTSRKAQEAAMLQLEOKATKQELITYYINK 173
Qy 121 YMSNGNTGMOTAAONTYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRRLVL 180
Db 174 YMSNGNTGMOTAAONTYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRRLVL 233
Qy 181 SEMKNOGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNLYLKEVINOVEEETGYNLLT 240
Db 234 SEMKNOGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNLYLKEVINOVEEETGYNLLT 293
Qy 241 TGMDEVYTNVDEAOKHMDIYNTDEVYVAPDELOVASTIVDSNGKYIAOLGARHOSNN 300
Db 294 TGMDEVYTNVDEAOKHMDIYNTDEVYVAPDELOVASTIVDSNGKYIAOLGARHOSNN 353
Qy 301 VSEGINQAVETNRDGMSTMKPTTDYAPALEYGVYDSTATIVHDEBYNPGTNPVYNNMDR 360
Db 354 VSEGINQAVETNRDGMSTMKPTTDYAPALEYGVYDSTATIVHDEBYNPGTNPVYNNMDR 413
Qy 361 GYFGITITQYALQOSRNPAYETLLNKVGLNRKPTFLNGLGIDYPSIHSNATSSMTTSSD 420
Db 414 GYFGITITQYALQOSRNPAYETLLNKVGLNRKPTFLNGLGIDYPSIHSNATSSMTTSSD 473
Qy 421 KKGASSEKMAAAYAFANGGTYYKPMYIHKYVFSDGSEKESNNGTAAKMTTAYMMTD 480
Db 474 KKGASSEKMAAAYAFANGGTYYKPMYIHKYVFSDGSEKESNNGTAAKMTTAYMMTD 533
Qy 481 MKKTYLVTGTRNAYLANLPQAGKTGTSNYTDEETENHIKTSQFVAPDELEFAGYTRKYSM 540
Db 534 MKKTYLVTGTRNAYLANLPQAGKTGTSNYTDEETENHIKTSQFVAPDELEFAGYTRKYSM 593
Qy 541 AVMTGYSRLPLVNGNGLTVAAKYVRSMMTYISEGSNPEDNMIIPGLRNGEFVKKNGAR 600
Db 594 AVMTGYSRLPLVNGNGLTVAAKYVRSMMTYISEGSNPEDNMIIPGLRNGEFVKKNGAR 653
Qy 601 STWNSPAPQPPSTRESSSSSDSTSOSSTPTTNNNTTTPNNNTTTPDOONON 660

Db 654 STWSSPAPQPPSTRESSSSSDSTSOSSTPTTNNNTTTPNNNTTTPDOONON 713
Qy 661 POPAOP 666
Db 714 POPAOP 719

RESULT 7
US-09-815-242-13423

; Sequence 13423, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13423

Query Match 99.5%; Score 3466; DB 22; Length 719;
Best Local Similarity 99.4%; Pred. No. 1.2e-287;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KIYDKRKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLENN 60
Db 54 KIYDKRKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLENN 113
Qy 61 LOSNSLOGSSTLTQOLIKLTYFSTSTSDQTSRKAQEAAMLQLEOKATKQELITYYINK 120
Db 114 LOSNSLOGSALTQOLIKLTYFSTSTSDQTSRKAQEAAMLQLEOKATKQELITYYINK 173
Qy 121 YMSNGNTGMOTAAONTYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRRLVL 180
Db 174 YMSNGNTGMOTAAONTYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRRLVL 233
Qy 181 SEMKNOGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNLYLKEVINOVEEETGYNLLT 240
Db 234 SEMKNOGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNLYLKEVINOVEEETGYNLLT 293
Qy 241 TGMDEVYTNVDEAOKHMDIYNTDEVYVAPDELOVASTIVDSNGKYIAOLGARHOSNN 300
Db 294 TGMDEVYTNVDEAOKHMDIYNTDEVYVAPDELOVASTIVDSNGKYIAOLGARHOSNN 353
Qy 301 VSEGINQAVETNRDGMSTMKPTTDYAPALEYGVYDSTATIVHDEBYNPGTNPVYNNMDR 360


```

Db 354 VSEGINQAVEINRDMGSTMKPTIDYAPALEGYVESTATIVHDEPNYGTNPVYNMNR 413
Qy 361 GYFGNITLLOYALQOOSRNVAVETLNKVGILNRAKTFPLNGCIDIPSHYSAISSNTTESD 420
Db 414 GYFGNITLLOYALQOOSRNVAVETLNKVGILNRAKTFPLNGCIDIPSHYSAISSNTTESD 473
Qy 421 KKYGASSEKMAAAYAAPANGCTYKPMYIHKVYFSDGSEKEFSNVGTRAMKETTAYAMTD 480
Db 474 KKYGASSEKMAAAYAAPANGCTYKPMYIHKVYFSDGSEKEFSNVGTRAMKETTAYAMTD 533
Qy 481 MKKTIVLYTGGRNAYLAMLPOAGKTGTSNYTDEIENHIKTSQFVAPDELFAGYTRKYSM 540
Db 534 MKKTIVLYTGGRNAYLAMLPOAGKTGTSNYTDEIENHIKTSQFVAPDELFAGYTRKYSM 593
Qy 541 AVWTGYSNRLTPLYGNGLYVAAKVRYSMTYLSGSPNPEDMNIPBGLYRNGEYFVRNGAR 600
Db 594 AVWTGYSNRLTPLYGNGLYVAAKVRYSMTYLSGSPNPEDMNIPBGLYRNGEYFVRNGAR 653
Qy 601 STWNSPAPQOPPESTESSSSSDSTSSSTTPSTNNSTTTNNNTTQOOSNTTPOOONON 660
Db 654 STWNSPAPQOPPESTESSSSSDSTSSSTTPSTNNSTTTNNNTTQOOSNTTPOOONON 713
Qy 661 POPAP 666
Db 714 POPAP 719

```

RESULT 8

```

US-10-072-851-13423
; Sequence 13423, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423
; LENGTH: 719
; TYPE: PRF
; ORGANISM: Streptococcus pneumoniae
US-10-072-851-13423

```

Query Match 99.5%; Score 3466; DB 24; Length 719;
 Best Local Similarity 99.4%; Pred. No. 1.2e-287;
 Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 KIYDNKNQIADIAGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
Db 54 KIYDNKNQIADIAGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
Qy 61 LOSNSILOGSSTLTQOLIKLTYFSTSDQTSISKRAQEWALAIQLEOKATKOELLTYINK 120
Db 114 LOSNSILOGSSTLTQOLIKLTYFSTSDQTSISKRAQEWALAIQLEOKATKOELLTYINK 173

```

```

Qy 121 VYKSNNGYGMQTAQNYVYGKDLNNLSLPOLALLAGHPQAPNOYDPSHPEAQRRLVYL 180
Db 174 VYKSNNGYGMQTAQNYVYGKDLNNLSLPOLALLAGHPQAPNOYDPSHPEAQRRLVYL 233
Qy 181 SEKNNGYISAEOYERAVNTPITDGLQSLKSASNPAYADNLYKEYINQVEETGYNLT 240
Db 234 SEKNNGYISAEOYERAVNTPITDGLQSLKSASNPAYADNLYKEYINQVEETGYNLT 293
Qy 241 TGMDEVYTNDOEAKHLMIDYNTDEVYAAPDDELQVASTIVDSNCKVIAOLGARQSSN 300
Db 294 TGMDEVYTNDOEAKHLMIDYNTDEVYAAPDDELQVASTIVDSNCKVIAOLGARQSSN 353
Qy 301 VSEGINQAVEINRDMGSTMKPTIDYAPALEGYVSDSTATIVHDEPNYGTNPVYNMNR 360
Db 354 VSEGINQAVEINRDMGSTMKPTIDYAPALEGYVESTATIVHDEPNYGTNPVYNMNR 413
Qy 361 GYFGNITLLOYALQOOSRNVAVETLNKVGILNRAKTFPLNGCIDIPSHYSAISSNTTESD 420
Db 414 GYFGNITLLOYALQOOSRNVAVETLNKVGILNRAKTFPLNGCIDIPSHYSAISSNTTESD 473
Qy 421 KKYGASSEKMAAAYAAPANGCTYKPMYIHKVYFSDGSEKEFSNVGTRAMKETTAYAMTD 480
Db 474 KKYGASSEKMAAAYAAPANGCTYKPMYIHKVYFSDGSEKEFSNVGTRAMKETTAYAMTD 533
Qy 481 MKKTIVLYTGGRNAYLAMLPOAGKTGTSNYTDEIENHIKTSQFVAPDELFAGYTRKYSM 540
Db 534 MKKTIVLYTGGRNAYLAMLPOAGKTGTSNYTDEIENHIKTSQFVAPDELFAGYTRKYSM 593
Qy 541 AVWTGYSNRLTPLYGNGLYVAAKVRYSMTYLSGSPNPEDMNIPBGLYRNGEYFVRNGAR 600
Db 594 AVWTGYSNRLTPLYGNGLYVAAKVRYSMTYLSGSPNPEDMNIPBGLYRNGEYFVRNGAR 653
Qy 601 STWNSPAPQOPPESTESSSSSDSTSSSTTPSTNNSTTTNNNTTQOOSNTTPOOONON 660
Db 654 STWNSPAPQOPPESTESSSSSDSTSSSTTPSTNNSTTTNNNTTQOOSNTTPOOONON 713
Qy 661 POPAP 666
Db 714 POPAP 719

```

RESULT 9

```

US-09-107-532-5667
; Sequence 5667, Application US/09107532
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECIUM FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7308
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085598
; FILING DATE: May 14, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:

```



```
Query Match          50.3%; Score 1751; DB 22; Length 778;
Best Local Similarity 51.4%; Pred. No. 2.4e-140;
Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;

QY 1 KYDKNKLIADLGSERRVNAQANDIPDVLKAIYSIEDHREDFHGDITRILGAFLRN 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 81 KYIDINNEIFEDLGAKEKRELIDPNVLPOLKDAIVSVEDERRFYKHIGVDPIRIGSALS 140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 LOSNLSOGSSTLTQOLIKLTFFSTSTSDOTISRKAQEAAMLAIQLEKATKOEILYYINK 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 141 VKNGLGOGSSTLTQOLIKLTFFSTSTSDOTISRKAQEAAMAVRLERESKEBILYYINK 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 VYMSNGNTGMOTAAQNYGKDLNNLSLPOLALLAGMPAOPNOYDPYSHEPAADRRLVL 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 201 VYMAFGFEMETAEENYKHLSELDPOTALLAGMPAOPNSYDPTKPTDKERRDVYL 260
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 SEMKNOGYSABOYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINQVEBETGYNLT 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 261 YTMVDNKRKISKAEEKAKATPIDEGLVPLKASDNRKVVADNVYKEVINBKAKTKGKNYT 320
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 TGMDEVYTNVDOAKHLMIDYNTDEYVAYPDELOVASTIYDVNSGKYIAOLGARHOSN 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 321 DGLDITLMDMAKQKOLYDVNSDOYVAFPDCKMOVASTVIDVASGQVRAQIGRHIPDD 380
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 VSEFINGOAVETNRDMGSTMKPITDYAPALEYGYDSTATIVHDEPYNPGTNPYVAMDR 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 381 VOLGNLAVNTORDVGSYTKPIMDYGPALENLNT-STGLRLMWDKPKTKPGFTDIDVNSDL 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 GYFGNITLOVALQOSRNPVAVETLNKVGILNRAKTFNLGLGIDYPSIHYSNAISSNTESD 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 440 TYOGVITMRRAIMGSRNLTAVQTFDEVGKRNIMPFIKGLIDYKNLEASNAISSNTSDVD 499
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 -KKYGASSEKMAAAVAFANGCTYKPMYIHKVYFSDGSEKEFVNGVTRAMKETAAYMT 479
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 500 GDKIGISLKLAAVAAAFANNGIYKPKYYVKNVYENDGSTVDYQDGRKAMKDSIAYMT 559
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 480 DMKKTIVLYTGGRNAYLAMPQAGKTGTSNTDEIEENHIKTSQFVAPDELFGYTRKYS 539
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 560 DMLKDVNLNGTGFNGALIGLLOAAKTGTSNYTDEDLARMGTTEKGIADPSTFVGTTTHYA 619
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 540 MAVMTGYSNRLTPLYGNGLTVAAKYRSMNTYLSGSPNPEDMNIEGLYRNGEYFKKGA 599
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 620 VSWMTGYNDNRNPTIYQEYGYGASDVYREIMSYLSQVNSNDMDWQDVSVYRGVGNELYKDA 679
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 600 RSTWN-----SPAOPPSTESSSSSDSTOSSSTPTSTNNSTTTNNNNTOOSN 651
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 660 YEVPVQVLPSTTSSAPPESSSTYESSSTKEAESSSSSSSAPSSSEAPSTEQPASS 739
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 652 TTPDOONPOPAOP 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 740 SSAEQPATSEOPPER 754
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-10-072-851-10728
; Sequence 10728, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
```

```
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072.851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 10728
; LENGTH: 778
; TYPE: PRF
; ORGANISM: Enterococcus faecalis
; US-10-072-851-10728

Query Match          50.3%; Score 1751; DB 24; Length 778;
Best Local Similarity 51.4%; Pred. No. 2.4e-140;
Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;

QY 1 KYDKNKLIADLGSERRVNAQANDIPDVLKAIYSIEDHREDFHGDITRILGAFLRN 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 81 KYIDINNEIFEDLGAKEKRELIDPNVLPOLKDAIVSVEDERRFYKHIGVDPIRIGSALS 140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 LOSNLSOGSSTLTQOLIKLTFFSTSTSDOTISRKAQEAAMLAIQLEKATKOEILYYINK 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 141 VKNGLGOGSSTLTQOLIKLTFFSTSTSDOTISRKAQEAAMAVRLERESKEBILYYINK 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 VYMSNGNTGMOTAAQNYGKDLNNLSLPOLALLAGMPAOPNOYDPYSHEPAADRRLVL 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 201 VYMAFGFEMETAEENYKHLSELDPOTALLAGMPAOPNSYDPTKPTDKERRDVYL 260
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 SEMKNOGYSABOYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINQVEBETGYNLT 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 261 YTMVDNKRKISKAEEKAKATPIDEGLVPLKASDNRKVVADNVYKEVINBKAKTKGKNYT 320
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 TGMDEVYTNVDOAKHLMIDYNTDEYVAYPDELOVASTIYDVNSGKYIAOLGARHOSN 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 321 DGLDITLMDMAKQKOLYDVNSDOYVAFPDCKMOVASTVIDVASGQVRAQIGRHIPDD 380
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 VSEFINGOAVETNRDMGSTMKPITDYAPALEYGYDSTATIVHDEPYNPGTNPYVAMDR 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 381 VOLGNLAVNTORDVGSYTKPIMDYGPALENLNT-STGLRLMWDKPKTKPGFTDIDVNSDL 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 GYFGNITLOVALQOSRNPVAVETLNKVGILNRAKTFNLGLGIDYPSIHYSNAISSNTESD 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 440 TYOGVITMRRAIMGSRNLTAVQTFDEVGKRNIMPFIKGLIDYKNLEASNAISSNTSDVD 499
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 -KKYGASSEKMAAAVAFANGCTYKPMYIHKVYFSDGSEKEFVNGVTRAMKETAAYMT 479
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 500 GDKYGISLKLAAVAAAFANNGIYKPKYYVKNVYENDGSTVDYQDGRKAMKDSIAYMT 559
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 480 DMKKTIVLYTGGRNAYLAMPQAGKTGTSNTDEIEENHIKTSQFVAPDELFGYTRKYS 539
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 560 DMLKDVNLNGTGFNGALIGLLOAAKTGTSNYTDEDLARMGTTEKGIADPSTFVGTTTHYA 619
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 540 MAVMTGYSNRLTPLYGNGLTVAAKYRSMNTYLSGSPNPEDMNIEGLYRNGEYFKKGA 599
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 620 VSWMTGYNDNRNPTIYQEYGYGASDVYREIMSYLSQVNSNDMDWQDVSVYRGVGNELYKDA 679
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 600 RSTWN-----SPAOPPSTESSSSSDSTOSSSTPTSTNNSTTTNNNNTOOSN 651
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 660 YEVPVQVLPSTTSSAPPESSSTYESSSTKEAESSSSSSSAPSSSEAPSTEQPASS 739
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 652 TTPDOONPOPAOP 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 740 SSAEQPATSEOPPER 754
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-09-134-000-4939
; Sequence 4939, Application US/09134000A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; APPLICANT: Nucleic Acid and Amino Acid Sequences Relating to Enterococcus
```


; TITLE OF INVENTION: FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-005
 ; CURRENT APPLICATION NUMBER: US/09/134,000A
 ; CURRENT FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 6810
 ; SEQ ID NO 4939
 ; LENGTH: 789
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 ; US-09-134-000-4939

Query Match 50.3%; Score 1751; DB 15; Length 789;
 Best Local Similarity 51.4%; Pred. No. 2,4e-140;
 Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;

QY 1 KIYDNKNOLIADLGSERRVNAOANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 60
 DB 92 KIYDINNEIFEDLGAERKRLIOPNDVQLKDAIVSEDRRKYKHIGVPIRILIGSALS 151
 QY 61 LOSNSLOGGSTLTLOOLIKLTFESTSDOTISRKAQEWALAIQLEKATKOEILTYINK 120
 DB 152 VANGSIGSGSTLTLOOLIKLTFESTSDOTISRKAQEWALAIQLEKATKOEILTYINK 211
 QY 121 VYMSNGNYGMOTPAONYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAAODRRNLVL 180
 DB 212 VYMANGYEMETAEVNYGKHLSEDLPLQALLAGMPQAPNSYDPTKDTAKERDYL 271
 QY 181 SEMKNOGYISAOYEKAVNPTITDGLQSLKSASNPAYMDNLKEVINOVEETGYNLT 240
 DB 272 YMYDMKKISKAEYERKATPIDEGLVPLKASDNRKVVNDYVKEVINERKAKTGNVYT 331
 QY 241 TGMADVYTNVDOEAKHLMIDYNTDEVYAVPDELQVASTIVVSNKGYIAOLGARHOSN 300
 DB 332 DGLDITYNLDMNAQOLYIVNSDOYVAFPPDKMOYASTIVIVASQVQAOLGGRITPDD 391
 QY 301 VSEGINQAVETNRDWSITMKPTDYAPALEGYVDSTACIVHDEPNYDPTNTPVYVMDR 360
 DB 392 VOLGNLAVNTQORDVGSYVKKPIMDYGPALLENLNY-STGRILAVDKPTKYPGTDIDVNSDL 450
 QY 361 GYFNGTITLOALQOSNRVAVETLNKVGILNRAKTPILNGIGIYPSHYNSNAISSNTESD 420
 DB 451 TYOGVITMERRALMSRNTAVOTFDEVGKENIMPTKIGLIDYKNLEASNAISSNTSDVD 510
 QY 421 -KRYGASSSEKMAAAYAFANGCTYKPMYIHKYVFSDSGESEFSNNGTAMKETTAYMMT 479
 DB 511 GKRYGSSSLKLAAYAFANGCTYKPMYIHKYVFSDSGESEFSNNGTAMKETTAYMMT 570
 QY 480 DMKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEIEIHNKTSQFVAPDELPAGYTRKYS 539
 DB 571 DMLKDVLCNGTGFNGAIPGLIOAKGTGSNYTDEIEDLARMGTEKGIAPDSTFGVYTHYA 630
 QY 540 MAVWNGYSRRLTPLYGNGILTVAAKYRSMYTLSEGSNEDMNIPBGLYRNGEYVKNKA 599
 DB 631 VSWWNGYNDRNPTIYQYEGYIASDYRELMSYLSQVNSDDWVOPPSVYVAGNELVKA 690
 QY 600 RSTWN-----SPADOPSTSESSSSSSSTSSSTSTSTSTSTSTSTSTSTSTSTSTSTST 651
 DB 691 YEVPVVOVLPSTSSAPPESSSYESSSTKFAESSSSSSSESAPSSSEAPSTEDQPASS 750
 QY 652 TTPDOONOPPOAP 666
 DB 751 SSAEQPATSEQPEP 765

RESULT 15
 US-08-116-541-4
 ; Sequence 4, Application US/08116541
 ; GENERAL INFORMATION:
 ; APPLICANT: Measure, H. Robert
 ; APPLICANT: Pearce, Barbara J.
 ; APPLICANT: Toumanen, Elaine
 ; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND

; TITLE OF INVENTION: CELLULAR VACCINES BASED THERON
 ; NUMBER OF SEQUENCES: 45
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klaubert & Jackson
 ; STREET: 411 Hackensack Avenue
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/116,541
 ; FILING DATE: 19930901
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 600-1-069
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201 487-5800
 ; TELEFAX: 201 343-1684
 ; TELEX: 133521
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 320 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-116-541-4

Query Match 46.5%; Score 1621; DB 5; Length 320;
 Best Local Similarity 99.7%; Pred. No. 7,2e-130;
 Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIYDNKNOLIADLGSERRVNAOANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 60
 DB 5 KIYDINNEIFEDLGAERKRLIOPNDVQLKDAIVSEDRRKYKHIGVPIRILIGSALS 64
 QY 61 LOSNSLOGGSTLTLOOLIKLTFESTSDOTISRKAQEWALAIQLEKATKOEILTYINK 120
 DB 65 LOSNSLOGGSALTLOOLIKLTFESTSDOTISRKAQEWALAIQLEKATKOEILTYINK 124
 QY 121 VYMSNGNYGMOTPAONYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAAODRRNLVL 180
 DB 125 VYMSNGNYGMOTPAONYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAAODRRNLVL 184
 QY 181 SEMKNOGYISAOYEKAVNPTITDGLQSLKSASNPAYMDNLKEVINOVEETGYNLT 240
 DB 185 SEMKNOGYISAOYEKAVNPTITDGLQSLKSASNPAYMDNLKEVINOVEETGYNLT 244
 QY 241 TGMADVYTNVDOEAKHLMIDYNTDEVYAVPDELQVASTIVVSNKGYIAOLGARHOSN 300
 DB 245 TGMADVYTNVDOEAKHLMIDYNTDEVYAVPDELQVASTIVVSNKGYIAOLGARHOSN 304
 QY 301 VSEGINQAVETNRDWS 316
 DB 305 VSEGINQAVETNRDWS 320

Search completed: June 13, 2002, 08:44:42
 Job time: 248 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:40:34 ; Search time 20.69 Seconds
(without alignments)
2575.142 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484
Sequence: 1 KIYDNKNQIADLGSERRVY.....TQSNTPDQGNPQAP 666

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 244551 seqs, 79999522 residues

Total number of hits satisfying chosen parameters: 244551

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCU_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3484	100.0	666	4 US-08-961-083-2	Sequence 2, Appl1
2	529.5	15.2	821	5 US-09-769-744A-122	Sequence 122, App
3	526	15.1	789	5 US-09-540-209B-8941	Sequence 8941, Ap
4	384	11.0	790	5 US-09-540-209B-9358	Sequence 9358, Ap
5	162	4.6	875	5 US-09-914-543-6	Sequence 6, Appl1
6	160.5	4.6	2893	5 US-09-882-227-522	Sequence 522, App
7	159	4.6	615	5 US-09-882-227-408	Sequence 408, App
8	153.5	4.4	930	5 US-09-200-650D-3	Sequence 3, Appl1
9	148	4.2	188	5 US-09-540-209B-6592	Sequence 6592, Ap
10	147.5	4.2	595	5 US-09-673-605A-34	Sequence 34, Appl1
11	146	4.1	2843	5 US-09-442-489B-2	Sequence 2, Appl1
12	143	4.1	1010	5 US-09-914-543-10	Sequence 10, Appl1
13	141	4.0	2842	5 US-09-442-489B-7	Sequence 7, Appl1
14	141	4.0	2843	5 US-09-987-482-1	Sequence 1, Appl1
15	139.5	4.0	706	7 US-60-360-039-1865	Sequence 1865, Ap
16	134	3.8	994	5 US-09-935-625-7547	Sequence 7547, Ap
17	134	3.8	994	5 US-09-935-625-28205	Sequence 28205, A
18	134	3.8	1006	5 US-09-935-625-7546	Sequence 7546, Ap
19	134	3.8	1006	5 US-09-935-625-28204	Sequence 28204, A
20	133.5	3.8	478	5 US-09-445-289B-34	Sequence 34, Appl1
21	133.5	3.8	541	7 US-60-360-039-9976	Sequence 9976, Ap
22	133.5	3.8	627	7 US-60-360-039-18489	Sequence 18489, A
23	132	3.8	1014	7 US-60-360-039-1608	Sequence 1608, Ap
24	131.5	3.8	1084	7 US-60-360-039-11155	Sequence 11155, A
25	131.5	3.8	1518	7 US-60-360-039-22243	Sequence 22243, A
26	131.5	3.8	10431	1 PCT-US02-11734-310	Sequence 310, App

27	131	3.8	989	5 US-09-935-625-7548	Sequence 7548, Ap
28	131	3.8	989	5 US-09-935-625-28206	Sequence 28206, A
29	131	3.8	1031	5 US-09-935-625-8784	Sequence 8784, Ap
30	131	3.8	1031	5 US-09-935-625-29529	Sequence 29529, A
31	131	3.8	1043	5 US-09-935-625-8783	Sequence 8783, Ap
32	131	3.8	1043	5 US-09-935-625-29528	Sequence 29528, A
33	129	3.7	708	5 US-09-540-209B-6575	Sequence 6575, Ap
34	128	3.7	1029	5 US-09-540-209B-10169	Sequence 10169, A
35	127.5	3.7	1609	7 US-60-360-039-1535	Sequence 1535, Ap
36	125	3.6	655	7 US-60-360-039-4713	Sequence 4713, Ap
37	125	3.6	655	7 US-60-360-039-7473	Sequence 7473, Ap
38	125	3.6	1166	5 US-09-200-650D-7	Sequence 7, Appl1
39	124.5	3.6	1169	5 US-60-360-039-22288	Sequence 22288, A
40	124	3.6	650	7 US-60-360-039-22177	Sequence 22177, A
41	124	3.6	889	7 US-60-360-039-14305	Sequence 14305, A
42	124	3.6	889	7 US-60-360-039-14653	Sequence 14653, A
43	124	3.6	889	7 US-60-360-039-15126	Sequence 15126, A
44	123	3.5	1504	7 US-60-360-039-2471	Sequence 2471, Ap
45	122	3.5	648	5 US-09-540-209B-9400	Sequence 9400, Ap

ALIGNMENTS

RESULT 1
US-08-961-083-2
Sequence 2, Application US/08961083
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: P3340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-961-083-2
Query Match 100.0%; Score 3484; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 1.7e-217;
Matches 666; Conservative 0; Mismatches 0; Gaps 0;
OY 1 KIYDNKNQIADLGSERRVNAQNDIPDVLKAIYSDHRFPDRGIDTRIIGAFLEARN 60


```

Db      1 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHREFDHDGIDTIRLGAFLRN 60
QY      61 LOSNSLOGSSTLTOOLIKLTYFSTSDOTISRKAQEAAMILOLEOKATKOELLTYINK 120
Db      61 LOSNSLOGSSTLTOOLIKLTYFSTSDOTISRKAQEAAMILOLEOKATKOELLTYINK 120
QY      121 VYMSNGYMGQTAQNYKYKDLNLSLPLALLAGMPQAPNOYDPYSHPBAADRNLVL 180
Db      121 VYMSNGYMGQTAQNYKYKDLNLSLPLALLAGMPQAPNOYDPYSHPBAADRNLVL 180
QY      181 SEMKNGYISAEQYKAVNPTIDGLQSLKASNPAYADNLIKETINOVEETGNNLT 240
Db      181 SEMKNGYISAEQYKAVNPTIDGLQSLKASNPAYADNLIKETINOVEETGNNLT 240
QY      241 TCGMDVYNTVDQAKHLMQIYNTDEYVAYPDELOVASTIVDSNGKVIAOLGARHQS 300
Db      241 TCGMDVYNTVDQAKHLMQIYNTDEYVAYPDELOVASTIVDSNGKVIAOLGARHQS 300
QY      301 VSEGINQAVETNRDMSSTMKPTIDYAPALEYGYDSTATIVHDEPNYDGTNTPVYNNMR 360
Db      301 VSEGINQAVETNRDMSSTMKPTIDYAPALEYGYDSTATIVHDEPNYDGTNTPVYNNMR 360
QY      361 GFPGNTLTOYALQOSRNVAVEITLKNVGLNRAKTFNLGIDIPSIHYSNAISSNTESD 420
Db      361 GFPGNTLTOYALQOSRNVAVEITLKNVGLNRAKTFNLGIDIPSIHYSNAISSNTESD 420
QY      421 KRYGASSEKMAAAYAFANGGTYYKPMYIHKVYFSDGSEKEFSNVGTRAKETATAMMD 480
Db      421 KRYGASSEKMAAAYAFANGGTYYKPMYIHKVYFSDGSEKEFSNVGTRAKETATAMMD 480
QY      481 MKKTVLYTGGRNAYLAMLPOAGKTGTSNYTDEIEINHITKTSQFVAPDELFAGYTRKYS 540
Db      481 MKKTVLYTGGRNAYLAMLPOAGKTGTSNYTDEIEINHITKTSQFVAPDELFAGYTRKYS 540
QY      541 AWTGSGNLTPLVGNGLVAAKYVRSMTYLSGSGNPEDMNIPDLIYRNGEYFNGAR 600
Db      541 AWTGSGNLTPLVGNGLVAAKYVRSMTYLSGSGNPEDMNIPDLIYRNGEYFNGAR 600
QY      601 STWNSAPQOPPESTESSSDSSTSTPSTNNSTTTNNNTTQOQNTTPOOQON 660
Db      601 STWNSAPQOPPESTESSSDSSTSTPSTNNSTTTNNNTTQOQNTTPOOQON 660
QY      661 POPAP 666
Db      661 POPAP 666

```

RESULT 2
US-09-769-744A-122
Sequence 122, Application US/09769744A

```

; GENERAL INFORMATION:
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hamill, Sean B
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21122WO
; CURRENT APPLICATION NUMBER: US/09/769,744A
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/GB99/02452
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: GB 981636.3
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125329
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 122
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-744A-122

```

```

Query Match      15.2%; Score 529.5; DB 5; Length 821;
Best Local Similarity 27.2%; Pred. No. 2,4e-26;
Matches 203; Conservative 115; Mismatches 270; Indels 159; Gaps 32;

QY      9 LTAOLGE--RRVNAQANDIPTDLVKAIVSIEDHREFDHDGIDTIRLGAFLRNLOS 63
Db      115 VIASIESDILRTSISEQISENKAITATEDHEHFKHGVKAVIRATTLKRFV--GLGS 173
QY      64 NSLOGSSTLTOOLIKLTYFSTSDOTISRKAQEAAMILOLEOKATKOELLTYINKYVM 123
Db      174 SS--GGSTLTOOLIKQOYVDA---PTLARKAAEYDALALERAMKDELITTYLVNAPF 228
QY      124 SNGNYGM-----QTAQNYKYKDLNLSLPLALLAGMPQAPNOYDPYSHP----- 169
Db      229 GRNNKQNTAGARQAAGIEFGVDASQUTVPQAALFLAGLPSPITPYSPEYNTGELKSEDL 288
QY      170 EAAODRRNLVLEMKKQGYISAEQYKAVNPTI-----TDGLQSL 209
Db      289 EIGLRRAKAVLSMYRTGALSKEDEYSQYKDYLDKODFLPSGTVTGISRDIYFTTLAEQ 348
QY      210 KSASNPAYMDNY-LKEVINQVE-----BETGYNLTGMDVYNTVDQAKH 256
Db      349 ERMYDILAOQDNVSAKELNKAETQKFRDLAKKEIENGCKITTT-----IDCKIISA 401
QY      257 LMDIYNTDEVAVYPPD---ELQVASTIVDSNGKVIAOLGARHQSNSVSEGINQAVETNR 313
Db      402 MQSA--VADGYLLDDGTFGVEGVNLMDNQGAILGFGGRNYQDNQ---NHADTKR 456
QY      314 DMSGTRKPTIDYAPALEYGYDSTATIVHDEPNYDGTNTPVYNNMRGTFGNTLTOYALO 373
Db      457 SPASTTKPLAAGIAIDQGLMS--ETILSNYPFNFGNPMIMANSKG--TGMMTLGELN 514
QY      374 QSRNPAVEET---LNKVNRAKTFNLGIDIPSIHYSNAISSNTESDCKRYGASSEK 430
Db      515 YSNINIRAYMTYRMLREKQVD--VKGVMKKGYELP-----EGIEISLPM 556
QY      431 AA-----AYAFANGGTYYKPMYIHKVYFSDGS--EKEFSNVGTRAKETAYMM 478
Db      557 GGGIEVTAQHTNGYQTLANNNGYHOKHVISKEADGKVVEYQDKPQVYSKATATIM 616
QY      479 TDMKTVLYTG-----TGRNAYLAMLPOAGKTGTSNYTDEIEINHITKTSQFVAP 529
Db      617 QGLLREVLSSRVTTTEKSNLTSLNPTLANADWIGKTGTTN--QDENWMLSLPRLTLGG- 674
QY      530 LFAGYTRKYSMAVWGTGYSN-----RLTPLV--GN--GLTVAAYKVSMTYLS 574
Db      675 -WIGHDHNHSLSRAGYSNNNSNMALVNAIQOASPSIMGNERFALDPVYKSEV--LKST 732
QY      575 GSNPEDMNIP-----EGLYRNGEYFVKNGARST-----WNSPAPQOP- 611
Db      733 GKPEKVSVEGKEVEVTGVTSTYMWKSGAPATYSFAIGSGDADYQNMWSSIVGSLPT 792
QY      612 PSTESSSSSDSSTSSSTSTPSTNNSTTTNNNTTQOQNTTPOOQON 668
Db      793 PSSSSSSSSSDSSTSSSTPSTSSRA 819

```

RESULT 3

```

US-09-540-209B-8941
Sequence 8941, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 8941
; LENGTH: 789
; TYPE: PRT

```


Matches	119; Conservative	71; Mismatches	196; Indels	178; Gaps	26;
OY	215	YPAVDNLIKETINOVEETGYNLITLGM-D-VYTNVDEQAQKHLMDIYITDEYVAPDDE	273		
Db	200	YAATLITLHAKKLOMPADSTKG---TTGTDIYVSNCTGAQG---FYTSYGYV-YDE	249		
OY	274	LQVASTYDVNSGKYALQAGARHSSNVSFFGINQAVETNRDMSGTMRKPIITDVPAPALEYGV	333		
Db	250	LAMGALMLMRATGEGDFYLEQAQKHYYGLMGFE-NQITTPPYTMSLGM---NDNA-----YAV	301		
OY	334	YDSTATIVHDEPYNT-----PGTNP-----VYNNDRGYE-----GNITLOYAL	372		
Db	302	YVLMALVALGDEVYHADAQRYLIDHWSVGEGRNTPNGILIVDSGVNRYRANAGYLAIFYAD	361		
OY	373	QOSRNVPAVETLANKYGLNA-----KFFLNGLGIDYR-SIH-----YSMAISS	414		
Db	362	ALGSHPLDYDRYHNFGRKQIDHILCDNDPNQSYVVGFGNFPINVHHRSGHSWSISIN	421		
OY	415	NTTESDKTKYGASSEKMAAAYAFANGGYT---KPYIHKHYVESQSEKEFSNVGIRAAK	471		
Db	422	PYNQNRHVLGY-----AVAGGPGQDGYIEEDRNDYVQNEVATD-----	458		
OY	472	ETTAYVMTDMKTVLYLTGTG-RNAYLAMLPAQKGTGTSNYTDEE-----	514		
Db	459	-----YNSGFTSAVALYLDHYGAPLANFPPEPESEYELVGAKINSGG	502		
OY	515	-----IENHITSQVAVDELFACTRTKRYSAAVWTGYS-NRLTPLYG-----NG	557		
Db	503	NRFVEKAVIQRH-ETTPAQGRKDLRYMRIFYDLSEVFAAGYSLNDLTVASGYNQASDVNG	561		
OY	558	LT-----VAAKYVRSM-----TYLSEGSNPEDWN-----IPEGL	587		
Db	562	LQHMGNVYVVAQRYDDVFPFGGSAHREYQFRSLPTTINLMAEMDTNDRPSPSLV	621		
OY	588	YRNGEFVF-----KNGARSTWSPADQPPSTESSSSSSSDSTSGSSTTPSTN	636		
Db	622	TYVDSLSLYGIDAPKIPLYDANGL-LMGEEPRPGGTSSSSSSSSSSSSSSSSSSSSSSSS	679		
OY	637	NSTTTNPNNNITQOSMTTPQONON	660		
Db	680	SS	703		

RESULT 6
US-09-883-227-522
; Sequence 522, Application US/09882227
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois

```

APPLICANT: Odomen, Raymond P.
TITLE OF INVENTION: Identification of Polynucleotides
TITLE OF INVENTION: Encoding Novel Helicobacter Polypeptides in the Helicobacter
TITLE OF INVENTION: Genome
FILE REFERENCE: 06132/047002
CURRENT APPLICATION NUMBER: US/09/882,227
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 08/902,615
PRIOR FILING DATE: 1997-07-29
NUMBER OF SEQ ID NOS: 638
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 522
LENGTH: 2893
TYPE: PRT
ORGANISM: Helicobacter pylori
US-09-882-227-522

Query Match          4.6%; Score 160.5; DB 5; Length 2893;
Best Local Similarity 19.5%; Pred. No. 0.096;
Matches 142; Conservative 103; Mismatches 251; Indels 233; Gaps 35;

```

QY	63	SNSLGGSGTLLFOOLILLYFEFSTSTQDTLSRKQAQWMLAIQLEOKPT----	KEOILTYEYI	118
Db	251	STSEFGTITMATN-----TAFSNAAGSFT-----	FEENATFSQAKKNGGTYTF	292
QY	119	NKVSNGNGYGMQTAQANYGRDL-----	NNLSLPOLALLAGMPAQPNQYD	164
Db	293	NKEFSATNTTAFSSGSFNKGVSSFNCGTSFNSASNYTFDQATFQNSSFNCGTTFPNNQTN		352
QY	165	PYS---HPEAADRRLVLJSEKKNQGYISAEQYERAVNTPITDGLQSLKSASNYPRAYMDN		221
Db	353	PNNMAQHPDIONSSEFSGNATLTK--GEVNFQQAFFNNSHQLT-----	IONASFNNATFN	405
QY	222	YLKEVINQVEEETGYLLTGMQDVYTNVQEAQKHLMDIYNTEDEYVAYPPDELOVASTIV		281
Db	406	TGKITI---EKASAFNNTTFNITSVDTN-----	NMSYTGCVLLSGNDLKNSTL-	451
QY	282	DVSNCRV-----IAOLGAR-----HQSNSVF-----	GINOAVETNRD--	314
Db	452	DHGSKITLTAQGTFFMLTSLGSEKSYTILNSSGCIITYSNLNLAINGLTBSALTNESLSN		511
QY	315	-----W-----	GSTAKPTDYAPALEGYVDSTAITVHDEPNYPG	350
Db	512	POSAQGLMDIITYNGVITQGLNENATSKP--TDSSPSKS----	STNSIQVYGVGYKIGD	566
QY	351	TMTPRYNNMRCGYFEGNTTLQYALQOQSRNPAVELNKNVGLNRAKTPFLNGJIDVPSHYSN		410
Db	567	T---YKLOETFSHNSIITQALESGTYTP-----	PPVYNGSKFD--	ISASN 607
QY	411	AISNNTTESDCKYKYGASSEKMAAAYAAFAANGTYKRP-----	MYIH--KYVFS--DGS--	458
Db	608	YNADMPWYDHXYIIPKSO-----	NFTESGYLLPDSVQIWSGYTNSFKQTFSGANOSNLV	661
QY	459	-----EKESNVGSTRAMKETTYATMTDMMKTUULTYTGGRNAYLAMLPOAGKTGTSNY		510
Db	662	IGYNSTWTDHANYSSSGTYSFGDTS-----	GSALNGHCGPWPYOCTGTJNG	707
QY	511	TOBEIENHITKTSOFVAPDELPAGYTRKYSMAVWTGYSNRLTPTLVNGL-----	TVA-	561
Db	708	TYSAHVYI--TANLNSGNIGTG-----	GAAN--	LIFNGVDSINJANATITQ 751
QY	562	--AKVYRSMKTYLSEB--SNPEDMNIPEGLYRN-----	GEFEVK-----	596
Db	752	HNAGIYSSMTFSTQGMDSQNLN--GLNSNCKLSYGVGTFPTEAKDKGFIINAGAYF		808
QY	597	-----NGARSTWNPAPQOPPTSESSSSSDSSTSTPSTNNSTTTNPNNNTOQSN		651
Db	809	EXTNENFGSGYQSCDILNFSNNNQFNSGFSPEISAKNAFNNANFNNSASFNFNNSNATTS		868
QY	652	TTPDOONON 660		
Db	869	EYGDFTTN 877		

```

RESULT 7
US-09-882-227-408
; Sequence 408, Application US/09882227
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oocmen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding Novel Helicobacter
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/047002
; CURRENT APPLICATION NUMBER: US/09/882,227
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/902,615
; PRIOR FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 408

```


[illegible]

```

RESULT 12
US-09-914-543-10
: Sequence 10, Application US/09914543
: GENERAL INFORMATION:
: APPLICANT: DIVERSA CORPORATION
: APPLICANT: LAM, David
: APPLICANT: MATHUR, Eric
: TITLE OF INVENTION: ENDOGLUCANASES
: FILE REFERENCE: DIVER1150-5
: CURRENT APPLICATION NUMBER: US/09/914,543
: CURRENT FILING DATE: 1998-11-22
: PRIOR APPLICATION NUMBER: PCT/US97/08793
: PRIOR FILING DATE: 1997-05-22
: PRIOR APPLICATION NUMBER: US 08/651,572
: PRIOR FILING DATE: 1996-05-22
: NUMBER OF SEQ ID NOS: 52
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 10
: LENGTH: 1010
: TYPE: PR1
: ORGANISM: Tereidinbacter
US-09-914-543-10

```

Query Match 4.1%; Score 143; DB 5; Length 1010;
 Best Local Similarity 19.5%; Pred. No. 0.31;
 Matches 131; Conservative 111; Mismatches 146; Gaps 29;

[illegible]

```

Db      265 DGRPAVNTEQGMDFIKONNTHSLNMSVSDKLBASLTVQ-----PG--PIISGM-----312
QY      364 GNTLQVALQOOSRNVPAVETLNK-VGINRAKTEFLNGIGIDYPSIIHSMASISNTTSDKK 422
Db      313 -----ASDLTASGTLFKRKIVSNMCTTIGNG-----SSSSSSSSSSSSSSSSSS 355
QY      423 YGASEKMAAAYAAFPANGCTYYKRMYYIHKVFS---DGSEKFSNVCTRAMKETAYMMMT 479
Db      366 SSSSSSSSSSSSSSSSGSTGGACAGVNNYPMWTADMSGAYNHANMADQWVONSILYRAN 415
QY      480 DMKATVLYGTGRNAYLAWLPOAKTKTSTNYTDEIEINHAKTSIOFVAPDELFAGYTRKYS 539
Db      416 KVTNIV-----PGSDASWTSILGACGCGNGSTISSSSSSSSSSSSSSSSSTGGGSSSSSS 471
QY      540 MAWVTGYSNRLLTPLVNGNLTVAAKRVYSMNTYISEGSPEDMWTPEGLYRNGEYFRKG- 598
Db      472 SSSSSSSSSSSSSSGGCGCTEVCWMYGGQYPLICNNTHSGCMWENNOICIGRQCEGONG 531
QY      599 -----ARSTNNSPAPQOPPTRESSSSSDSTSSSTTSTNNSTTTNNNNTOOSNTT 653
Db      532 AGGVVSNCTGSGTSSS--SSSSSSSSSSSSSSSSSSSGTSGTSSSTSSSSSSSSSSSSST 588
QY      654 PDQGNQNPQAPQ 666
Db      589 GSSGMPGPRVNDP 601

```

```

RESULT 13
; Sequence 7, Application US/09442489B
; GENERAL INFORMATION:
; APPLICANT: Albertsen, Hans
; APPLICANT: Anand, Rakesh
; APPLICANT: Carlson, Mary
; APPLICANT: Groden, Joanna
; APPLICANT: Hedge, Philip John
; APPLICANT: Josiyy, Geoff
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Markham, Alexander Fred
; APPLICANT: Nakamura, Yusuke
; APPLICANT: Thliveris, Andrew
; APPLICANT: Vogelstein, Bert
; APPLICANT: White, Raymond L.
; TITLE OF INVENTION: APC Antibodies
; FILE REFERENCE: 001107.78817
; CURRENT APPLICATION NUMBER: US/09/442,489B
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: US 08/452,654
; PRIOR FILING DATE: 1995-05-25
; PRIOR APPLICATION NUMBER: US 08/289,548
; PRIOR FILING DATE: 1994-08-12
; PRIOR APPLICATION NUMBER: US 07/741,940
; PRIOR FILING DATE: 1991-08-08
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2842
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-442-489B-7

```

Query Match 4.0%; Score 141; DB 5; Length 2842;
Best Local Similarity 18.7%; Pred. No. 1.8;
Matches 129; Conservative 99; Mismatches 213; Indels 250; Gaps 31;

```

Oy  56  AFRRLJONS-----LOGSFTLQOLKLTFTSTSDOTISRKAEAMHAILOEQ 106
      | | | | |
Db   723  AALRLMLNAPRAKTKDANINMPSGSLPSLR-----KQ 756
      | | | | |
Oy  107  KATCOELITTYINKVYNSNGY-----GMQTAONYEYKDLNINSLPOLALLAGMPOA 159
      | | | | |
Db   757  KALEHEIDLAOHLSFTFNDINDINSLKASHRSKORHKOBLDY-----799
      | | | | |

```



```

0Y 160 PNQDVPYHPREAADRRN----LYLSMKKQGIISAQYERKAVNPPIIDGLOSLKASANY 215
Db 800 ---FDJNHDNRSDNFGNTMTLSPY-----LNTTV-----LPSSSSS 836
0Y 216 PAYMNPYKFEVINQVEEETGYNLLTGTGMQDYTVWVQEAQKHLMDIYNTEDEYVAPDDE-- 273
Db 847 RGSIDSSSEKDRLEBERG-----IOLGVY-----HREATNP 869
0Y 274 -----LQVASTIVDVNSNGKVIQAOLGARHQSSNVSFGINQAVETNRDWCST--MKPIITD 324
Db 870 GTSSKRGLOISTFAAQIA--KVMEVSAIHTS-----QEDRSSGSTTELCVTD 916
0Y 325 YAPALEYVYNSTATIYHDEPYNPGNINPYVMNDGFCGNITLOYA--LQSRNPAVET 383
Db 917 ERNALR-----RSSAAHHSMTNF-----TKSENSNR-----TCSMPAKLEYRRS--SDDS 962
0Y 384 LNKVGLN-----RAKTFUNGGLID-----YPS-----HYNSAISRNTTESDK 421
Db 963 LNSVSSSGGYCKRQMKPSTIESYSEDDSEKFCYGYQYPADLAHKIHSANHMDNDNGELDT 1022
0Y 422 -----KYGASSEKMAAAYAFANGTYTKPMYIHKVFEPSDSEKEKESNVGTRAMKETTA 475
Db 1023 PLYNSLKY--SDEQLNSGRQSPQSGNEWAPKHIIDEKIKQSRQSRN-----QSTT 1073
0Y 476 YMAIDMAKTVLTYTGNRAIILAMPQAGKTGTNTYDELEENHIKTSCVAVADELFAGYT 535
Db 1074 YPV-----YTESDDKHLKFOPHFGQOEVCSPYR 1102
0Y 536 RKYSMAVWTGYSNRLPFLVGNGLTLVAAKYRSMW-----PLYSGNSPEDMNPIE 585
Db 1103 SR-----GANGETNRVGSNHCINONVSQSLQOEDDIYEDDKPTNYSKRYSEEQHEE 1155
0Y 586 GLYRNGEVEFVKNGARSTWNSP-----AFQOPSTRESSSSSDSTOSSTT--PSTNN 637
Db 1156 ERPTNYSIKYNEEKRAHV--DQPIDYSLKAYNDIPSSQKQSFSPSKSSGQSSKTEHMSSS 1214
0Y 638 STTTPNPNNTQOQN-----TTPDOONOPQA 664
Db 1215 ENTSTPSSNAKRONQLPSSAQSRSGQPOKA 1245

```

```

RESULT 14
US-09-987-482-1
; Sequence 1, Application US/09987482
; GENERAL INFORMATION:
; APPLICANT: BHANDARI, POONAM
; APPLICANT: SHASHIDHARA, L.S.
; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF
; TITLE OF INVENTION: DRUGS AND OTHER SUBSTANCES
; FILE REFERENCE: 056859-0134
; CURRENT APPLICATION NUMBER: US/09/987,482
; CURRENT FILING DATE: 2002-03-21
; NUMBER OF SEQ. ID NOS.: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ. ID NO. 1
; LENGTH: 2843
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-482-1

```

Query Match	4.0%	Score 141	DB 5	Length 2843
Best Local Similarity	18.7%	Pred. No. 1.8		
Matches 129	Conservative 99	Mismatches 213	Indels 250	Gaps 31

```

Oy  56  AFRRLJONS-----LOGGFTLQQLKLYFSTNSDQYISKRADEAMLALDQ 106
      |||||:|
Db  724  ALRLMLANRPRAKTKDANIMSPGSSLPYLVR-----KO 757
      |||||:|
Oy  107  KATKOELLTYINKVYMSGNY-----GMQTAQNYGKDLNLSLPOLLALAMPQA 159
      |||||:|
Db  758  KALELDELAOHLSEFTFDINDIMSPKASHRSQRRKOSLYGYV-----800
      |||||:|

```

QY	160	PNQDYPHYHPAAADRRN----	LYLSBMKNOGYSIAQYERAVNTPITDGLQSLKSAASY	215
Db	801	---EPTNRHNDNRSDNENTGNMTVLSPY	-----LMTTV-----LPSSSSS	837
QY	216	PAYMDNYLKEVINQVEEETGYNLLTGTGMQVYTVNQEOAKHLMADIYNTDEYVAPDDE--	273	
Db	838	RGSIDSSSEMDRLEBERG-----	IGLGNV-----HATENP	870
QY	274	-----LQVASTIVDVSNGKVIQAOLGARHQSNSVSFGINQAVETNRDWCST--MKPTD	324	
Db	871	GTSSKRGLOISTAAQIA--KVMEVSAIHTS-----	QEDRSSGSTTELCVTD	917
QY	325	YAPALEYVYDSTATYIHDEPYNPGCNTPYVYMDQGYCIGNTLOYA--LQOSRNPANET	383	
Db	918	ERNALR-----RSSAAHTHSNTYF-----TKSENSNR--TCSMPAKLETYRS--SDDS	963	
QY	384	LNKYCLN-----RAKTFPLNGCID-----YPS-----HYNSAISNTTESDK	421	
Db	964	LNSVSSSGCYCKRQMKPJSLESTSEDDSEKFCYGCYQYPADLANKHHSANHMDNDNGELDT	1023	
QY	422	-----KYGASSEMAAAYAFANGGTYYKPMYIHKVPESDSEKEFSNVGTRAKETTA	475	
Db	1024	PINYSLKX--SDEQLNGSRQSPQSNEMAPKHIIEDELIKQESQNRN-----QSTT	1074	
QY	476	YMDMDMAKTYLTYTGTGRNATLAMPQAGKTGTSNYTDELEHNIKITSOVAPDELFACTY	535	
Db	1075	YPV-----	-----YTESDDKHLKFOHPFGQOEVSYPYR	1103
QY	536	RKYSMAWTCYSNRLPYLGNGLTVAAKYVRSMM-----	-----YLSGSPPECDMNP	585
Db	1104	SR-----GANGSETNRKVSNGHICINONYSOLCOEDDYEDKPPINYSEBRSSEEOHDEE	1156	
QY	586	GLYRNGEEVFENKGRSTYMNPP-----APOQPTRESSSSSDSTOSSSSTT--PSTNN	637	
Db	1157	ERPNTYSIKYEAKRHV--DQPIDYSLKVAADIPSSQKQFSFSKSSSGSKSXTHEMSSSS	1215	
QY	638	STTTPNNNTQOSN-----TTPDOONOPQA	664	
Db	1216	ENTSTPSSNAKRONQLHPSSAOSNSGQPOKA	1246	

```

RESULT 15
US-60-360-039-1865
; Sequence 1865, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1865
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-60-360-039-1865

```

Query Match	4.0%;	Score 139.5;	DB 7,	Length 706;
Best Local Similarity	19.0%;	Pred. No. 0.32;		
Matches 146;	Conservative 115;	Mismatches 227;	Indels 279;	Gaps 40.

```
07      58 LRN-LQNSLQGGSLTLOOLIKLYFFSTIS-----DGTISK-----AQ 96  
        ||| ||::||| :| :|||: |::| |:  
Db     3 LRNIQAASSLLSGLSAAD-----SSSTTGDDGYAPSIIPCPBDDTLVRAASGLSTAE 55  
  
07     97 EAMLAIQLEKATYQOEIIITYINRVYSNGNYGMOCTAONYGKDNLNNLSLPDI----- 150
```



```
Db 56 TDMK---KRDAYTKEALHSEFLSR---ATSNFSDTSLSTLFSS---NSSNPKIGIACSG 107
OY 151 ----ALLAGMPQAPNOYDPSHPEAODRNLVLEMKNOGYSAEQYEKAVN----- 199
Db 108 GGYRAMLGAGACI-----RANDRIDGANEHIGLGLOSSTYISGLSGGNMLTG 156
OY 200 TPITDGLQSLKASNYPAYMD---NYLKEVINQVEETGYNLLTGMVYTNVDQEAQ-- 254
Db 157 TLAMNNMTSVQEIYDHMESEDSINMTITSIVN----PGSNLYTITIERMESIYQEVQAK 211
OY 255 -----KHLMD---IYNTEYVAYPDDELQVA-STIVDY--SNGKV--IAQLGAR 295
Db 212 SPAGFNISLIDLAMARALSYNF--FPLPDAGSALTWSSLRDVPKNGEMPLPIIVADOR 269
OY 296 HOSSNVSGINQAVETNMDGSTMKPIIDYAPALFYGYDSTATIVHDEPYNPGTNT-- 353
Db 270 YPGTIV-INLNAT-----LFEFTF-FEMGSWDPSLNAFTD--VKYIGTNVTN 312
OY 354 --PYVNMDRGYEGNITLQYALQOOSHNV-----PAVETLNKVGILNRAKTFNLGLGIDYP 404
Db 313 GKPV-NKQOCVSGYDNAGFVITASLNFESLEASTSTYKMINSPANKYVNNLSQDD 371
OY 405 SI-----HYSNA----- 411
Db 372 DIAIYAANPFKDFEVDNRNYSIVDADDLFLVGDGEDGQNLPLVPLIKKERDLVFPAL 431
OY 412 -ISSNTTES-----DKKYGASSEKMAAAYA---AFANGTYKKPMYT----- 449
Db 432 DISDNTDESMPGVCMTNTYERQYKQKGMAFPYPVPVNTFNLGLTNKPTFFGCDAKN 491
OY 450 -----HKVFSDDSEKEFS-NVGTIRAMKETTAYMMTDMKTVLTYGTGRNAYLAWLP 500
Db 492 LTDLEIPIPLVYIIPNTKHSFENGOSTLKN---YNVTERL-----GMIRNGF----- 536
OY 501 QAGTGTGNSNYTDE-----IENHIKTSQFVAPDELFAGYTRKYSMAVWTGYSNRL 550
Db 537 EAATMG--NFTDASNFLGICGCAIRRKQESLNATLPREC---TKCFADYCWNG----- 585
OY 551 TPLVNGGLTVAKAYRSMATYILSEGSNPEDMNIPGCLIRNGEFVFNKNGARSTWNSPAPQ 610
Db 586 -----TLSTANPE-----LSGNSTYOSGAIASAISEATDG 616
OY 611 PP-----STESSSSSDSTSQSSSTPTNNSTTTNPNNTQOQN 651
Db 617 IPTALLGSSSTSGNTTSNTSTSTSSNVTNSNSSSNTTLNNSSSSS 663
```

Search completed: June 13, 2002, 08:41:39
Job time: 65 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OK protein - protein search, using sw model

Run on: June 13, 2002, 08:41:59 ; Search time 34.55 Seconds
(without alignments)
2141.104 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 666
Sequence: 1 KIDNRKNTLADLSESRVN.....TQSSNTTPDQONONQPAQP 666

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 10

Total number of hits satisfying chosen parameters: 30

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A.GeneSeq_032802:*

```
1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	666	19	AAW55063
2	330	49.5	682	17	AAW04359
3	330	49.5	719	22	AAU37830
4	245	36.8	320	16	AAU70153
5	206	30.9	420	20	AAV56106
6	14	2.1	778	22	AAU35135
7	11	1.7	266	19	AAV86004
8	11	1.7	821	21	AAV81757
9	10	1.5	110	19	AAW55113
10	10	1.5	119	22	AAW69173
11	11	1.5	284	20	AAV49227

12	10	1.5	284	20	AAV32100
13	10	1.5	406	20	AAV49225
14	10	1.5	406	20	AAV32098
15	10	1.5	428	20	AAV49238
16	10	1.5	428	20	AAV32110
17	10	1.5	438	22	AAV92949
18	10	1.5	446	20	AAV49140
19	10	1.5	446	20	AAV49143
20	10	1.5	459	20	AAV49251
21	10	1.5	459	20	AAV32190
22	10	1.5	460	20	AAV49250
23	10	1.5	460	20	AAV32189
24	10	1.5	462	22	AAU03646
25	10	1.5	655	20	AAV49226
26	10	1.5	655	20	AAV32099
27	10	1.5	694	21	AAV81653
28	10	1.5	774	22	AAU36453
29	10	1.5	781	22	AAU35684
30	10	1.5	1704	15	AAV49657

ALIGNMENTS

```
RESULT 1
AAW55063
ID AAW55063 standard: Protein; 666 AA.
XX
AC AAW55063:
XX
DE 02-OCT-1998 (first entry)
XX
DE Streptococcus pneumoniae SP001 protein.
XX
KW Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
KW detection; pneumonia; otitis media; meningitis.
XX
OS Streptococcus pneumoniae.
XX
PN W09818930-A2.
XX
PD 07-MAY-1998.
XX
PF 30-OCT-1997; 97WO-US19422.
XX
PR 31-OCT-1996; 96US-0029960.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Chol GH, Hromocky J A, Johnson LS, Kunsch CA;
XX
DR WPI: 1998-272224/24.
XX
DR N-PSDB: AAV27323.
XX
PT Nucleic acid encoding antigenic peptide(s) from Streptococcus
PT pneumoniae - or their epitope-containing fragments, useful in
PT protective or therapeutic vaccines, and for diagnosis
XX
PS Claim 11; Page 48; 118pp; English.
XX
XX
The present sequence represents a protein from Streptococcus pneumoniae.
The nucleic acid sequence encoding the streptococcus pneumoniae protein
can be useful in vaccines for inducing protective antibodies against
Streptococcus pneumoniae, for treatment or prevention of infection e.g.
pneumonia, otitis media or meningitis. Probes based on the nucleic acid
are used to detect Streptococcus infection (by usual hybridisation or
amplification methods), also for isolating Streptococcus genes or their
allelic variants. The protein can be used similarly to detect specific
antibodies in standard immunoassays, especially for diagnosing or
monitoring infections. Antibodies which bind the protein are used to
detect corresponding antigens, to purify the protein and for passive
immunisation (optionally coupled to a toxin). Vaccines are administered,
e.g. by injection, orally or through the skin, typically at 0.01-1000
```



```

|||||
Db 508 maylawpqaqkqtsnytleeeienhiktsqfvpadelifgyrkyrkyssmawtqysnrlt 567
OY 552 PLVNGSLTVAAKVYRSMWTYLSGSGNPEDWNIPEGLYRNGEFVKNGARSTW 603
Db 568 plvngsltvaakvyrsmwtlylsgsgnpedwnipegllyrngelvfkngrstcw 619

RESULT 3
AAU37830
ID AAU37830 standard; Protein: 719 AA.
XX
AC AAU37830;
XX
DT 14-FEB-2002 (first entry)
XX
DE Streptococcus pneumoniae cellular proliferation protein #259.
XX
KM Antisense; prokaryotic cellular proliferation protein;
KM antibiotic; antibacterial; drug design.
XX
OS Streptococcus pneumoniae.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Hesselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS55689.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 13423; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 719 AA;

```

```

Query Match 49.5%; Score 330; DB 22; Length 719;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 530; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 72 LTOGLIKITFVSTSTSDOTISRKQAQEWLAQLQKAKTKOELIYYIKKYMSNNGYQMO 131
Db 125 ltqglkltfyststsdqtliskqeaawlaqlqekatkqelllyllylnkkymsngygmq 184
OY 132 TAAQNYGKGDNNLSLPDLALACMPQAPNOYDPSHPEDADRNLVSEKKNNGYISA 191
Db 185 taqnyy9kdlmnlslpqlallagmpapnqydpshpeaaqdrnlvlsenkngyisa 244
OY 192 EOYEKAVNPITTDGLQSLKASANYPAYMDNLYKEVINOVEEBTGYNLLTGGMDVYNTNDQ 251
Db 245 egyekavnpittdqglqslksasnypayndnylkevinoveetgynlltggmdvylndq 304
OY 252 EAQKHLWDIYTFDEVAVPDELOVASTIVDSNGKVTIAQLGARHQSSNVSGINQAVET 311
Db 305 eaqkhlwdiyltdevaypddelqvasltvdvsnqkvlaqlgarhqsnsvsgfngavet 364
OY 312 NRDKSTPMKPRITDYAPALEYGYOSTATIVHDEPNYPGTNPVYNNMGRGYFGNTLQYA 371
Db 365 nrwstmkpiltidyapaleyyvestatlvndepynpgtntpvyndrgrfgnltlqya 424
OY 372 LQGRNVPVAVETLNKVGILNRAKTFNLGLIDYPSIHYSNATSSNTTESDKRYGASSEKMA 431
Db 425 lqgrnvpavetlnkvglnraktfnlglidypsihylnalssnteedkrygassekma 484
OY 432 AAYAFANGSTYKPRYTHKVVPSDSEKFEPSNNGTRAMKETTYAMTDMKTYLTYGTG 491
Db 485 aayaafangstykprmythkvvpsdsekefepsnngtramketlaymtdmktvlysgtg 544
OY 492 RNAYLAWLPQAGKGTSTNYTDEIEENHKTQSOFVAPDELFGYTRKYSMAVWTGYSNRLT 551
Db 545 rmaylawlpqaqkqtsnytleeeienhiktsqfvpadelifgyrkyrkyssmawtqysnrlt 604
OY 552 PLVNGSLTVAAKVYRSMWTYLSGSGNPEDWNIPEGLYRNGEFVKNGARSTW 603
Db 605 plvngsltvaakvyrsmwtlylsgsgnpedwnipegllyrngelvfkngrstcw 656

RESULT 4
AAR70153
ID AAR70153 standard; Protein: 320 AA.
XX
AC AAR70153;
XX
DT 14-FEB-1996 (first entry)
XX
DE Streptococcus pneumoniae strain SPRU42 Exp2.
XX
KW Exp2; export protein; pbp1a; p1pa; exp1; exp3; pad1;
KW virulence determinant; permease like protein;
KW penicillin binding protein 1a; pyruvate oxidase; regulatory element;
KW acellular vaccine; antibody.
XX
OS Streptococcus pneumoniae.
XX
FN WO9506732-A2.
XX
PD 09-MAR-1995.
XX
PF 01-SEP-1994; 94WO-US09942.
XX
PR 01-SEP-1993; 93US-0116541.
PR 18-MAY-1994; 94US-0245511.
XX
PA (UVRQ ) UNIV ROCKEFELLER.
XX
PI Masure HR, Pearce BJ, Tuomanen E;
XX
DR WPI; 1995-115448/15.
DR N-PSDB; AAO83241.

```


XX Novel gene fragments encoding specific bacterial exported proteins
PT - specifically of *S. pneumoniae*, useful as vaccines
XX
PS Claim 35; Page 88-9; 168pp; English.
XX This sequence represents exp2. The DNA encoding this sequence is
CC identical to that for pona which encodes penicillin-binding protein 1A
CC (Pbp1a). This sequence is involved in adhesion of bacteria to target
CC cells. This sequence is an exported protein of *S. pneumoniae*. Export
CC proteins are the proteins in pathogenic bacteria that are virulence
CC determinants. Other export proteins include p1pa (see AAF70152), exp1,
CC exp3, and pad1 (encoded by the sequence shown in AAF83259). This
CC sequence can be inserted into an expression vector (preferably a
CC bacterial expression vector) to provide for high levels of expression of
CC the protein. The protein can then be used in the production of an
CC acellular vaccine. These vaccines are used to provide protection from
CC Gram positive bacterial infection. Antibodies against export proteins
CC can be used for diagnosis of infection and in passive immune therapy.
XX
SQ Sequence 320 AA:
Query Match 36.8%; Score 245; DB 16; Length 320;
Best Local Similarity 100.0%; Pred. No. 1e-233;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 72 LTGOLIKLITFTSTSPDITSRKAOEAWLAIOLKATKOELLITTYINKVYMSNGYGMQ 131
Db 76 lTgqLiKlItYtStSdqlStKqaEaWlAlqLeqatKqellTtYInKvYmsNgYgmQ 135
OY 132 TAAQNYTGGKDLNNLSLPOLALAGMPQAPNQDPYSHPEAODRRNLVSEKNGYISA 191
Db 136 taaqnyTgGkdlnnlSpolAlaGmpqApnQdPyShPeaOdrrnlVseKngyIsa 195
OY 132 EOYEKAVNTPITDGLQSLKASNTPAYMDNLTKEVINOVEETGYNLLTTGMDYTTNDQ 251
Db 136 eQyEkAvnTpItDglQslKasNtPaYmDnLtKeVInOvEeTgYnLLtTgmDyTtnDq 255
OY 252 EAKQKLMIDYNDDEVYAVPDDLOVASTVDVNSGVINOLGARHSSNVSGINQAVET 311
Db 256 eaqKkLmIdYndDevYavPddLoVastVdVnSGvInoLgarHssNvsgInqAvet 315
OY 312 NRDMG 316
Db 316 nrDmg 320
RESULT 5
AAV56106
ID AAV56106 standard; Protein: 420 AA.
XX
AC AAV56106;
XX
DT 07-FEB-2000 (first entry)
XX
DE Streptococcus pneumoniae pbp1a TER isolate a) protein sequence.
XX
KM Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1a;
KW transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
XX detection; identification; pneumococcal meningitis.
XX
OS Streptococcus pneumoniae.
XX
PM ZA9807024-A.
XX
PD 28-APR-1999.
XX
PF 05-AUG-1998; 98ZA-0007024.
XX
PR 01-AUG-1997; 97ZA-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.

PA (UWI-) UNIV WITWATERSRAND.
XX (MEDI-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
XX
DR WPI, 1999-601770/51.
XX
DR N-PSDB; AA235939.
XX
PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
PS useful for the diagnosis of pneumococcal meningitis
XX
PS Claim 11; Fig 4; 63pp; English.
XX
XX A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1a gene. The products and methods can be used for detecting
CC *S. pneumoniae*, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant *S. pneumoniae* strains in a sample.
CC The methods can be used for detecting *S. pneumoniae* strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC *S. pneumoniae* with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1a transpeptidase encoding region (TER) isolate protein
CC sequence from the present invention.
XX
SQ Sequence 420 AA:
Query Match 30.9%; Score 206; DB 20; Length 420;
Best Local Similarity 99.5%; Pred. No. 5.4e-195;
Matches 406; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 196 KAVNPEITDGLQSLKASNTPAYMDNLTKEVINOVEETGYNLLTTGMDYTTNDQ 255
Db 5 kavnPeItDglQslKasNtPaYmDnLtKeVInOvEeTgYnLLtTgmDyTtnDq 255
OY 256 HLMIDYNDDEVYAVPDDLOVASTVDVNSGVINOLGARHSSNVSGINQAVETNRDM 315
Db 65 hlmIdYndDevYavPddLoVastVdVnSGvInoLgarHssNvsgInqAvetnrDw 124
OY 316 GSTMKRPTDYAPALEGVYDSTATVHDEPYNPGNFPYVNMDBRFGNITLOVALQS 375
Db 125 gStmKRpTdYapAlEgVydStaTvhDePynPgNfPyVnmDBrFgnItLoValQs 184
OY 376 RNVAVETLNKVLNRAKFTLNGLIDYPSIHYSNAISSNTTESDCKYGASSEKMAAAYA 435
Db 185 rnvaVeTlnKvLnraKfTlNgLIdYpSiHysNaIssntTesdckYgasseKmaaya 244
OY 436 AFANGGTYYKPMYIHKVPSDSGESEKFSNVGTRAKMETTAYMMTDMKTLVLTGTGRNAY 495
Db 245 aFangGtYyKpmYiHkVpsDsgEsEkfSNvgTramkettayMmtDmkTlVlsgtGrnay 304
OY 496 LAMLPQAGKTGTSNTYDELENIHKTSQFVAPDELFACTTRKYSMAVWVGYSNRLLTPVG 555
Db 305 lAmLpqaGktGtsNtYdeLeNiHkTsQfVapDeLfaCtTrkYsMaVwVgYSnrLLtpVg 364
OY 556 NGLTVAKVYRSMMYTLSEGSNPEDMNIPEGLYRNGEFPFKNGARSTW 603
Db 365 nglTvaKvYrsMmYtlSeGsnPeDmniPeGlyRngeFpfKngarStw 412
RESULT 6
AAU35135
ID AAU35135 standard; Protein: 778 AA.
XX
AC AAU35135;
XX
DT 13-FEB-2002 (first entry)


```

XX DE Enterococcus faecalis cellular proliferation protein #422.
XX KW Antisense; prokaryotic cellular proliferation protein;
XX KW antibiotic; antibacterial; drug design.
XX OS Enterococcus faecalis.
XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US09180.
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207727P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX DR WPI: 2001-611495/70.
XX DR N-PSDB; AAS52994.
XX PS New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids -
XX PS Example 3; Seq ID No 10728; 511pp; English.
XX PS The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the
XX CC genes, their use in the discovery of novel antibiotics, the essential
XX CC genes themselves and the encoded proteins. The prokaryotes used are
XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX CC invention is also useful for the identification of potential new targets
XX CC for antibiotic development. The antisense nucleic acids can also be used
XX CC to identify proteins used in proliferation, to express these proteins,
XX CC and to obtain antibodies capable of binding to the expressed proteins.
XX CC The proteins can be used to screen compounds in rational drug discovery
XX CC programmes. The antisense nucleic acid sequence is also useful to screen
XX CC for homologous nucleic acids which are required for cell proliferation in
XX CC a wide variety of organisms. The present sequence represents an
XX CC essential prokaryotic cellular proliferation protein.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 778 AA;

```

Query Match 2.1%; Score 14; DB 22; Length 778;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 66 LOGSTLQQLIKL 79
DB 146 1qg9stltqqlikl 159

```

RESULT 7
 AAY86004
 ID AAY86004 standard; Protein; 266 AA.
 AC AAY86004;
 XX

```

DT 10-APR-2000 (first entry)
XX DE S. pneumoniae derived protein #213.
XX KW Treatment; prevention; disease; diagnosis; gene therapy; screening;
XX KW bacterial; antimicrobial; antibiotic; pathogenesis; infection.
XX OS Streptococcus pneumoniae.
XX PN WO9806734-A1.
XX PD 19-FEB-1998.
XX PF 15-AUG-1997; 97WO-US14436.
XX PR 16-AUG-1996; 96US-0024022.
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
XX PI Stodola RK;
XX DR WPI: 1998-159452/14.
XX DR N-PSDB; AA296344.
XX PS Streptococcus pneumoniae proteins and related DNA - useful for
XX PT screening compounds for antibacterial activity
XX PS Claim 5; Page 493-494; 640pp; English.
XX CC This invention describes novel isolated Streptococcus pneumoniae
XX CC polynucleotides (see AAY86173-796494) and their encoded proteins (see
XX CC AAY85792-186182). The DNA, vectors and host cells described in the
XX CC method of the invention are useful for the recombinant expression of the
XX CC polypeptides. The polypeptides are useful for treatment or prevention of
XX CC disease, or diagnosis of disease related to expression or activity of
XX CC such a polypeptide. They can also be used to screen for compounds which
XX CC interact with and inhibit or activate such a polypeptide. The
XX CC polypeptides (or DNA encoding them, via gene therapy) are also useful
XX CC for inducing an immunological response in a mammal. The antagonists are
XX CC useful to inhibit such bacterial polypeptides. The polypeptides are
XX CC particularly useful to identify antimicrobial compounds and antibiotics.
XX CC They are also useful to determine their role in pathogenesis of
XX CC infection, dysfunction and disease.
XX SQ Sequence 266 AA;

```

Query Match 1.7%; Score 11; DB 19; Length 266;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 68 GGSYTLQQLIK 78
DB 107 ggsytlqqlik 117

```

RESULT 8
 AAY81757
 ID AAY81757 standard; Protein; 821 AA.
 AC AAY81757;
 XX 02-JUN-2000 (first entry)
 DE Streptococcus pneumoniae protein sequence ID113.
 DE Streptococcus pneumoniae infection; immunogen; antigen; diagnosis; AIDS;
 KW bacterial pneumonia; asplenia; heart disease; lung disease; alcoholism;
 KW kidney disease; diabetes; immunosuppressive disorder; otitis media;
 KW pneumococcal septicaemia; sinusitis; meningitis; therapy.
 OS Streptococcus pneumoniae.


```

XX PN WO200006738-A2.
XX XX
XX PD 10-FEB-2000.
XX XX
XX PF 27-JUL-1999; 99WO-GB02452.
XX XX
XX PR 27-JUL-1998; 98GB-0016336.
XX PR 19-MAR-1999; 99US-0125329.
XX XX
XX PA (MICR-) MICROBIAL TECHNIQS LTD.
XX PI Le Page RWF, Wells JM, Hanniffy SB, Hansbro PM;
XX XX
XX DR WPI: 2000-195301/17.
XX DR N-PSDB; AA291853.
XX XX
XX PT Streptococcal proteins and polynucleotides useful for diagnosis,
XX PT treatment and prophylaxis of bacterial infections
XX PS Claim 2: Page 59; 76pp; English.
XX XX
XX CC This sequence represents a Streptococcus pneumoniae protein of the
XX CC invention. The proteins (or their homologues, derivatives and/or
XX CC fragments) are useful as immunogens or antigens. Immunogenic or antigenic
XX CC compositions comprising the proteins are useful as vaccines and also in
XX CC diagnostic assays. The sequences are useful for the detection or
XX CC diagnosis of S. pneumoniae infection, by contacting a sample to be tested
XX CC with them. Agents capable of antagonising, inhibiting or interfering with
XX CC the function or expression of the protein or polypeptide are useful in
XX CC medical compositions in the treatment or prophylaxis of S. pneumoniae
XX CC infection. As the sequences can be used to treat S. pneumoniae infection,
XX CC they can be used to treat bacterial pneumonia, which has high rates in
XX CC young children, the elderly, and in patients with predisposing conditions
XX CC such as asplenia, heart, lung and kidney disease, diabetes, alcoholism,
XX CC or with immunosuppressive disorders, especially AIDS. They can also be
XX CC used to treat pneumococcal septicemia, otitis media, sinusitis, and
XX CC meningitis.
XX SQ Sequence 821 AA;

Query Match 1.7%; Score 11; DB 21; Length 821;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 GGSTLTQOLIK 78
DB 176 ggstltqqlik 186

RESULT 9
AAM55113
ID AAM55113 standard; Protein: 110 AA.
XX AC AAM55113;
XX XX
XX DT 02-OCT-1998 (first entry)
XX XX
XX DE Streptococcus pneumoniae SP0062 protein.
XX XX
XX KW Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
XX KW detection; pneumonia; otitis media; meningitis.
XX OS Streptococcus pneumoniae.
XX XX
XX PN WO9818930-A2.
XX PD 07-MAY-1998.
XX PF 30-OCT-1997; 97WO-US19422.
XX PR 31-OCT-1996; 96US-0029960.

```

```

XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Choi GH, Hromocky J A, Johnson LS, Kunsch CA;
XX XX
XX DR WPI: 1998-272224/24.
XX DR N-PSDB; AA27374.
XX XX
XX PT Nucleic acid encoding antigenic peptide(s) from Streptococcus
XX PT pneumoniae - or their epitope-containing fragments, useful in
XX PT protective or therapeutic vaccines, and for diagnosis
XX PS Claim 11; Page 70; 118pp; English.
XX XX
XX CC The present sequence represents a protein from Streptococcus pneumoniae.
XX CC The nucleic acid sequence encoding the Streptococcus pneumoniae protein
XX CC can be useful in vaccines for inducing protective antibodies against
XX CC Streptococcus pneumoniae, for treatment or prevention of infection e.g.
XX CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid
XX CC are used to detect Streptococcus infection (by usual hybridisation or
XX CC amplification methods), also for isolating Streptococcus genes or their
XX CC allelic variants. The protein can be used similarly to detect specific
XX CC antibodies in standard immunoassays, especially for diagnosing or
XX CC monitoring infections. Antibodies which bind the protein are used to
XX CC detect corresponding antigens, to purify the protein and for passive
XX CC immunisation (optionally coupled to a toxin). Vaccines are administered,
XX CC e.g. by injection, orally or through the skin, typically at 0.01-1000
XX CC (especially 10-300) mu g/ml per dose.
XX SQ Sequence 110 AA;

Query Match 1.5%; Score 10; DB 19; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
DB 18 sssssdst 27

RESULT 10
ABB69173
ID ABB69173 standard; Protein: 119 AA.
XX AC ABB69173;
XX XX
XX DT 26-MAR-2002 (first entry)
XX XX
XX DE Drosophila melanogaster polypeptide SEQ ID NO 34311.
XX XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX XX
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX XX
XX DR WPI: 2001-656860/75.
XX DR N-PSDB; ABL13276.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more

```


PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure: SEQ ID NO 34311; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL016175) and the encoded proteins
CC sequences (ABB57737-ABB72072).
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIRO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 119 AA;

Query Match 1.5%; Score 10; DB 22; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.17; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 617 SSSSSDSTS 626
Db 60 ssssdsts 69
|||||||

RESULT 11
AA49227
ID AA49227 standard; Protein: 284 AA.
XX
AC AA49227:
XX
DT 07-FEB-2000 (first entry)
XX
DE N-terminal truncate of Cbpa serotype type 4 polypeptide R1.
XX
KW Choline binding protein A; Cbpa; truncate; immune response; infection;
KW pneumococcal bacterium; vaccine.
XX
OS Streptococcus pneumoniae.
XX
PN WO9951187-A2.
XX
PD 14-OCT-1999.
XX
PF 07-APR-1999; 99WO-US07668.
XX
PR 07-APR-1998; 98US-0056019.
PR 07-APR-1998; 98US-0080878.
XX
PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI Tuomanen EI, Masure HR;
XX
DR WPI; 1999-620161/53.
DR N-PSDB; AA231403.
XX
PT Novel polypeptides, used to develop products for the diagnosis,
PT prevention and treatment of pneumococcal infections -
XX
PS Claim 3; Page 11; 85pp; English.
XX
CC The invention provides novel isolated polypeptides comprising the amino
CC acid sequence of an N-terminal choline binding protein A (Cbpa) truncate
CC (CbpaR). The polypeptides can be selected from sequences shown in
CC AA49225, AA49227, AA49230, AA49231, AA49233 and AA49238. The
CC polypeptides can be used for inducing an immune response in a subject
CC which has been exposed to or infected with a pneumococcal bacterium.
CC can also be used for preventing infection by a pneumococcal bacterium.
CC Vaccines comprising the polypeptides or encoding nucleic acids can be

CC used for treating a subject infected with or exposed to a pneumococcal
CC bacterium. Antibodies specifically binding the polypeptides can be used
CC for detection and diagnosis and for preventing pneumococcal attachment to
CC a mucosal surface. The products can be used in humans and other animals
CC such as domestic animals, such as feline or canine subjects, farm animals
CC such as bovine, equine, caprine, ovine, and porcine subjects, wild
CC animals (whether in the wild or in a zoological garden), research
CC animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats,
CC i.e. for veterinary medical use.
XX
SQ Sequence 284 AA;

Query Match 1.5%; Score 10; DB 20; Length 284;
Best Local Similarity 100.0%; Pred. No. 0.37; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 616 SSSSSDSTS 625
Db 113 ssssdsts 122
|||||||

RESULT 12
AA492100
ID AA492100 standard; Protein: 284 AA.
XX
AC AA492100:
XX
DT 01-FEB-2000 (first entry)
XX
DE Choline binding protein A (Cbpa) truncate R1.
XX
KW Choline binding protein; Cbpa; truncate; adhesion; immunogen;
KW vaccine; diagnosis; therapy; pneumococcus; sepsis; meningitis;
KW otitis media; pneumonia.
XX
OS Streptococcus pneumoniae.
XX
PN WO9951188-A2.
XX
PD 14-OCT-1999.
XX
PF 07-APR-1999; 99WO-US07669.
XX
PR 07-APR-1998; 98US-0056019.
PR 07-APR-1998; 98US-0080878.
XX
PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
PA (MEDI-) MEDIMUNE INC.
XX
PI Tuomanen EI, Masure HR, Wizemann TM, Johnson LS, Koenig S;
XX
DR WPI; 1999-633690/54.
DR N-PSDB; AA234453.
XX
PT New N-terminal choline binding protein A truncate polypeptides, used to
PT develop products for the diagnosis, prevention and treatment of
PT pneumococcal infections -
XX
PS Claim 3; Page 93-95; 160pp; English.
XX
CC The present sequence represents polypeptide R1, a new N-terminal
CC choline binding protein A (Cbpa) truncate derived from the Cbpa of
CC Streptococcus pneumoniae serotype type 4. Host-vector systems for
CC production of R1 and other Cbpa truncated polypeptides are provided.
CC The invention relates generally to novel N-terminal Cbpa truncates
CC and polynucleotides encoding them, host-vector systems, and
CC antibodies that specifically bind to the polypeptides. The
CC invention also relates to vaccines including Cbpa polypeptides,
CC which provide protection or elicit protective antibodies to
CC bacterial infection, specifically pneumococcus, and to use of
CC antibodies and antagonists against such polypeptides in diagnosis

XX C glutamicum protein fragment SEQ ID NO: 6703.
 XX
 XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KM organic acid synthesis.
 XX
 XX Corynebacterium glutamicum.
 OS
 XX
 XX EPI108790-A2.
 PN
 XX
 XX 20-JUN-2001.
 PD
 XX
 XX 18-DEC-2000; 2000EP-0127688.
 PF
 XX
 XX 16-DEC-1999; 99JP-0377484.
 PR
 XX 07-APR-2000; 2000JP-0159162.
 PR
 XX 03-AUG-2000; 2000JP-0280988.
 XX
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA
 XX
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 XX WPI: 2001-376931/40.
 DR N-PSDB; AAH68168.
 XX
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT
 XX
 XX Claim 17; SEQ ID NO: 6703; 246bp + Sequence Listing; English.
 PS
 XX
 XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 CC
 XX
 XX Sequence 438 AA;
 SQ
 Query Match 1.5%; Score 10; DB 22; Length 438;
 Best Local Similarity 100.0%; Pred. No. 0.55;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 616 SSSSSSDSST 625
 DB 69 sssssdsst 78
 ID AAY49140 standard; Protein; 446 AA.
 XX
 XX AAY49140;
 AC
 XX
 XX 17-JAN-2000 (first entry)
 DT
 XX
 XX Amino acid sequence of choline-binding protein fragment #4.
 DE
 XX
 XX Truncated surface binding protein; alpha helix; choline binding protein;
 KM vaccine; invasive bacterial infection; otitis media; sepsis;
 KM meningitis; lobar pneumonia infection; antibody; immature immune system;
 KM immunocompromised.

XX Streptococcus pneumoniae.
 OS
 XX
 XX Key Location/Qualifiers
 FH
 FT Misc-difference 171 /note= "Encoded by AAT"
 FT Misc-difference 172 /note= "Encoded by TAC"
 FT Misc-difference 173 /note= "Encoded by TTA"
 FT Misc-difference 174 /note= "Encoded by CAA"
 FT
 XX
 XX WO951266-A2.
 PN
 XX
 XX 14-OCT-1999.
 PD
 XX
 XX 06-APR-1999; 99WO-US07680.
 PF
 XX
 XX 07-APR-1998; 98US-0080878.
 PR
 XX 15-MAY-1998; 98US-0085743.
 XX
 XX (MEDT-) MEDIMUNE INC.
 PA
 XX
 XX Witzmann TM, Koenig S, Johnson LS;
 PI
 XX
 XX WPI: 1999-601465/51.
 DR N-PSDB; AAZ31080.
 XX
 XX New pneumococcal proteins useful as vaccines and for diagnosis of
 PT pneumococcal infections -
 PT
 XX
 XX Claim 10; Page 68-69; 98pp; English.
 PS
 XX
 XX AAY49137-Y49152 are amino acid sequences that are fragments of choline
 CC binding proteins (CBP). The fragments of the protein are the alpha helix
 CC forming parts of the CBPs from Streptococcus pneumoniae. The
 CC polypeptides do not contain the actual choline binding fragment. The
 CC polypeptides and the nucleotide sequences that encode them
 CC (AAZ31077-631092) are used in the invention, which relates to polypeptide
 CC truncates of a pneumococcal surface binding protein containing the highly
 CC conserved immunogenic alpha helical portion and no choline binding
 CC portion. The polypeptides are used as immunogens in a bacterial vaccine.
 CC The vaccine can be used for preventing (immunising) or treating invasive
 CC bacterial (especially pneumococcal) infections, especially otitis media
 CC (caused by S.pneumoniae), sepsis, meningitis and lobar pneumonia
 CC infections. Antibodies raised against the polypeptide are useful for
 CC detection, prevention (passive immunity) and treatment of S. pneumoniae
 CC infections. The vaccines are especially useful in immunocompromised
 CC patients, those with an immature immune system, or patients with an on
 CC going pneumococcal infection. The vaccine avoids unnecessary expense and
 CC provides broad protection against a range of pneumococcal serotypes and
 CC it produces an improved and enhanced effect in preventing bacterial
 CC infections.
 CC
 XX
 XX Sequence 446 AA;
 SQ
 Query Match 1.5%; Score 10; DB 20; Length 446;
 Best Local Similarity 100.0%; Pred. No. 0.56;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 616 SSSSSSDSST 625
 DB 114 sssssdsst 123
 ID AAY49143 standard; Protein; 446 AA.
 XX
 XX AAY49143;
 AC
 XX
 XX

DT 17-JAN-2000 (first entry)
XX
XX Amino acid sequence of choline-binding protein fragment #7.
DE
XX Truncated surface binding protein; alpha helix; choline binding protein;
KM vaccine; invasive bacterial infection; otitis media; sepsis;
KM meningitis; lobar pneumonia infection; antibody; immature immune system;
KM immunocompromised.
XX
OS Streptococcus pneumoniae.
XX
XX WO951266-A2.
XX
XX 14-OCT-1999.
XX
XX 06-APR-1999; 99WO-US07680.
XX
XX 07-APR-1998; 98US-0080878.
XX 15-MAY-1998; 98US-0085743.
XX
XX (MEDI-) MEDIMUNE INC.
XX
XX Wismann TM, Koenig S, Johnson LS;
XX WPI: 1999-601465/51.
XX DR N-PSDB; AA231083.
XX
XX New pneumococcal proteins useful as vaccines and for diagnosis of
PT pneumococcal infections -
XX
XX Claim 10; Page 72-73; 98pp; English.
XX
XX AA49137-149152 are amino acid sequences that are fragments of choline
CC binding proteins (CBP). The fragments of the protein are the alpha helix
CC forming parts of the CBPs from Streptococcus pneumoniae. The
CC polypeptides do not contain the actual choline binding fragment. The
CC polypeptides and the nucleotide sequences that encode them
CC (AA231077-231092) are used in the invention, which relates to polypeptide
CC truncates of a pneumococcal surface binding protein containing the highly
CC conserved immunogenic alpha helical portion and no choline binding
CC portion. The polypeptides are used as immunogens in a bacterial vaccine.
CC The vaccine can be used for preventing (immunising) or treating invasive
CC bacterial (especially pneumococcal) infections, especially otitis media
CC (caused by S.pneumoniae), sepsis, meningitis and lobar pneumonia
CC infections. Antibodies raised against the polypeptide are useful for
CC detection, prevention (passive immunity) and treatment of S. pneumoniae
CC infections. The vaccines are especially useful in immunocompromised
CC patients, those with an immature immune system, or patients with an on
CC going pneumococcal infection. The vaccine avoids unnecessary expense and
CC provides broad protection against a range of pneumococcal serotypes and
CC it produces an improved and enhanced effect in preventing bacterial
CC infections.
XX
XX Sequence 446 AA;
SQ

Query Match 1.5%; Score 10; DB 20; Length 446;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
DB 114 sssssdst 123

RESULT 20
ID AA49251
AA49251 standard; Protein; 459 AA.
XX
XX AA49251;
XX
XX 07-FEB-2000 (first entry)
XX

DE N-terminal region of CbpA polypeptide ATCC4.
XX
XX Choline binding protein A; CbpA; truncate; immune response; infection;
KM pneumococcal bacterium; vaccine.
KM
XX
OS Streptococcus pneumoniae.
XX
XX WO951187-A2.
XX
XX 14-OCT-1999.
XX
XX 07-APR-1999; 99WO-US07668.
XX
XX 07-APR-1998; 98US-0056019.
XX 07-APR-1998; 98US-0080878.
XX
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX Tuomanen EI, Masure HR;
XX WPI: 1999-620161/53.
XX
XX Novel polypeptides, used to develop products for the diagnosis,
PT prevention and treatment of pneumococcal infections -
XX
XX Disclosure; Fig 2A-B; 85pp; English.
XX
XX The invention provides novel isolated polypeptides comprising the amino
CC acid sequence of an N-terminal choline binding protein A (CbpA) truncate
CC (CbpA7). The polypeptides can be selected from sequences shown in
CC AA49225, AA49227, AA49230, AA49231, AA49233 and AA49238. The
CC polypeptides can be used for inducing an immune response in a subject
CC which has been exposed to or infected with a pneumococcal bacterium. They
CC can also be used for preventing infection by a pneumococcal bacterium.
CC Vaccines comprising the polypeptides or encoding nucleic acids can be
CC used for treating a subject infected with or exposed to a pneumococcal
CC bacterium. Antibodies specifically binding the polypeptides can be used
CC for detection and diagnosis and for preventing pneumococcal attachment to
CC a mucosal surface. The products can be used in humans and other animals
CC such as domestic animals, such as feline or canine subjects, farm animals
CC such as bovine, equine, caprine, ovine, and porcine subjects, wild
CC animals (whether in the wild or in a zoological garden), research
CC animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats,
CC i.e. for veterinary medical use.
XX
XX Sequence 459 AA;
SQ

Query Match 1.5%; Score 10; DB 20; Length 459;
Best Local Similarity 100.0%; Pred. No. 0.57;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
DB 127 sssssdst 136

RESULT 21
ID AA492190
AA492190 standard; Protein; 459 AA.
XX
XX AA492190;
XX
XX 01-FEB-2000 (first entry)
XX
XX N-terminal choline binding protein A (CbpA) truncate.
XX
XX Choline binding protein; CbpA; truncate; adhesin; immunogen;
KM vaccine; diagnosis; therapy; pneumococcus; sepsis; meningitis;
KM otitis media; pneumonia.
XX
XX Streptococcus pneumoniae.
OS Synthetic.
OS

XX WO9951188-A2.
 PN
 XX
 PT 14-OCT-1999.
 PT
 XX
 PF 07-APR-1999; 99WO-US07669.
 PR
 XX 07-APR-1998; 98US-0056019.
 PR 07-APR-1998; 98US-0080878.
 XX
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PA (MEDT-) MEDIMUNE INC.
 XX
 XX Tuomanen EI, Masure HR, Wizemann TM, Johnson LS, Koenig S;
 PI
 DR WPI; 1999-633690/54.
 XX
 PT New N-terminal choline binding protein A truncate polypeptides, used to
 PT develop products for the diagnosis, prevention and treatment of
 PT pneumococcal infections -
 PS
 XX Claim 47; Fig 2A-B; 160pp; English.
 CC
 CC The present sequence represents an N-terminal choline binding
 CC protein A (Cbpa) truncate, denoted ATCC4 Cbpa trun, derived from
 CC Streptococcus pneumoniae Cbpa. Claimed vaccines contain and
 CC N-terminal Cbpa truncate such as the present sequence, or a
 CC polypeptide comprising a conserved region of the Cbpa truncate.
 CC The vaccines provide protection or elicit protective antibodies to
 CC bacterial infection, specifically pneumococcus. Antibodies and
 CC and antagonists against the N-terminal Cbpa truncates are used in
 CC diagnosis and passive immunotherapy.
 CC
 SQ Sequence 459 AA;

Query Match 1.5%; Score 10; DB 20; Length 459;
 Best Local Similarity 100.0%; Pred. No. 0.57;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 616 SSSSSSDSST 625
 DB 127 sssssdsst 136

RESULT 22
 AAY49250
 ID AAY49250 standard; Protein; 460 AA.
 XX
 AC AAY49250;
 XX
 DT 07-FEB-2000 (first entry)
 DT
 XX
 DE N-terminal region of Cbpa polypeptide Ntype4.
 XX
 KW Choline binding protein A; Cbpa; truncate; immune response; infection;
 KW pneumococcal bacterium; vaccine.
 XX
 OS Streptococcus pneumoniae.
 OS
 PN WO9951187-A2.
 PN
 XX 14-OCT-1999.
 PD
 XX
 PF 07-APR-1999; 99WO-US07668.
 PF
 XX 07-APR-1998; 98US-0056019.
 PR 07-APR-1998; 98US-0080878.
 PR
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PA
 XX (MEDT-) MEDIMUNE INC.
 XX
 XX Tuomanen EI, Masure HR;
 PI
 PT

DR WPI; 1999-620161/53.
 XX
 XX Novel polypeptides, used to develop products for the diagnosis,
 PT prevention and treatment of pneumococcal infections -
 PT
 XX
 PS Disclosure; Fig 2A-B; 85pp; English.
 PS
 XX The invention provides novel isolated polypeptides comprising the amino
 CC acid sequence of an N-terminal choline binding protein A (Cbpa) truncate
 CC (CbpaT). The polypeptides can be selected from sequences shown in
 CC AAY49225, AAY49227, AAY49230, AAY49231, AAY49238 and AAY49238. The
 CC polypeptides can be used for inducing an immune response in a subject
 CC which has been exposed to or infected with a pneumococcal bacterium. They
 CC can also be used for preventing infection by a pneumococcal bacterium.
 CC Vaccines comprising the polypeptides or encoding nucleic acids can be
 CC used for treating a subject infected with or exposed to a pneumococcal
 CC bacterium. Antibodies specifically binding the polypeptides can be used
 CC for detection and diagnosis and for preventing pneumococcal attachment to
 CC a mucosal surface. The products can be used in humans and other animals
 CC such as bovine, equine, caprine, ovine, and porcine subjects, farm animals
 CC animals (whether in the wild or in a zoological garden), research
 CC animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats,
 CC i.e. for veterinary medical use.
 CC
 SQ Sequence 460 AA;

Query Match 1.5%; Score 10; DB 20; Length 460;
 Best Local Similarity 100.0%; Pred. No. 0.57;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 616 SSSSSSDSST 625
 DB 128 sssssdsst 137

RESULT 23
 AAY32189
 ID AAY32189 standard; Protein; 460 AA.
 XX
 AC AAY32189;
 XX
 DT 01-FEB-2000 (first entry)
 DT
 XX
 DE N-terminal choline binding protein A (Cbpa) truncate.
 XX
 KW Choline binding protein; Cbpa; truncate; adhesin; immunogen;
 KW vaccine; diagnosis; therapy; pneumococcus; sepsis; meningitis;
 KW otitis media; pneumonia.
 XX
 OS Streptococcus pneumoniae.
 OS Synthetic.
 XX
 PN WO9951188-A2.
 PN
 XX 14-OCT-1999.
 PD
 XX
 PF 07-APR-1999; 99WO-US07669.
 PF
 XX 07-APR-1998; 98US-0056019.
 PR 07-APR-1998; 98US-0080878.
 PR
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PA (MEDT-) MEDIMUNE INC.
 XX
 XX Tuomanen EI, Masure HR, Wizemann TM, Johnson LS, Koenig S;
 PI
 DR WPI; 1999-633690/54.
 DR
 XX
 PT New N-terminal choline binding protein A truncate polypeptides, used to
 PT develop products for the diagnosis, prevention and treatment of
 PT pneumococcal infections -

XX Claim 47; Fig 2A-B; 160pp; English.

PS

XX The present sequence represents an N-terminal choline binding

CC protein A (Cbpa) truncate, denoted Ntype 4 Cbpa trin, derived from

CC Streptococcus pneumoniae Cbpa. Claimed vaccines contain and

CC N-terminal Cbpa truncate such as the present sequence, or a

CC polypeptide comprising a conserved region of the Cbpa truncate.

CC The vaccines provide protection or elicit protective antibodies to

CC bacterial infection, specifically pneumococcus. Antibodies and

CC and antagonists against the N-terminal Cbpa truncates are used in

CC diagnosis and passive immunotherapy.

CC

XX

XX Sequence 460 AA:

QY 616 SSSSSSDSST 625

DB 128 sssssdst 137

Query Match 1.5%; Score 10; DB 20; Length 460;

Best Local Similarity 100.0%; Pred. No. 0.57;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OS

XX

XX RESULT 24

XX AAU03646

ID AU03646 standard; Protein; 462 AA.

XX

XX AAU03646;

XX

XX 12-SEP-2001 (first entry)

XX

XX Group B Streptococcus antigenic protein, ID-122.

XX

XX Group B Streptococcus; encapsulated bacterium; therapeutic; sepsis;

KW meningitis; neonate; antigenic; vaccine; infection; genital tract;

KW capsid polysaccharide vaccination.

XX

XX Streptococcus agalactiae.

XX

XX WO200132882-A2.

XX

XX 10-MAY-2001.

XX

XX 07-SEP-2000; 2000WO-GB03437.

XX

XX 07-SEP-1999; 99GB-0021125.

XX

XX (MICR-) MICROBIAL TECHNIQS LTD.

XX

XX Le Page RWF, Wells JM, Hanniffy SB;

XX

XX WPI: 2001-316444/33.

XX

XX N-PSDB; AAS07063.

XX

XX New polypeptides derived from Streptococcus agalactiae are useful to

PT provide detection of, and vaccination against, Group B Streptococcus

PT infections, particularly to prevent infection in neonates

XX

XX Claim 1; Fig 1; 178pp; English.

XX

XX AAU03601-AAU03722 represent Group B Streptococcus (Streptococcus

CC agalactiae) amino acid sequences of the invention. S. agalactiae is an

CC encapsulated bacterium which is a major pathogen of humans causing sepsis

CC and meningitis in neonates as well as adults. The S. agalactiae antigenic

CC polypeptides are used to vaccinate against Group B Streptococcus

CC infections, particularly to prevent infection in new born children

CC arising from the maternal genital tract. An immunogenic composition is

CC useful in the preparation of a medicament for the treatment or

CC prophylaxis of Group B Streptococcus infection. The invention does not

CC have the disadvantages of varied response rate associated with prior art

CC capsid polysaccharide vaccination against Group B Streptococcus.

XX

XX sequence 462 AA:

QY 313 RDMGSRMKPI 322

DB 369 rdwgrtmkpi 378

Query Match 1.5%; Score 10; DB 22; Length 462;

Best Local Similarity 100.0%; Pred. No. 0.58;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OS

XX

XX RESULT 25

XX AAU9226

ID AAU9226 standard; Protein; 655 AA.

XX

XX AAU9226;

XX

XX 07-FEB-2000 (first entry)

XX

XX Cbpa of serotype 4 amino acid sequence.

XX

XX Choline binding protein A; Cbpa; truncate; immune response; infection;

KW pneumococcal bacterium; vaccine.

XX

XX Streptococcus pneumoniae.

XX

XX WO951187-A2.

XX

XX 14-OCT-1999.

XX

XX 07-APR-1999; 99WO-US07668.

XX

XX 07-APR-1998; 98US-0056019.

XX

XX 07-APR-1998; 98US-0080878.

XX

XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX

XX Tuomanen EI, Masure HR;

XX

XX WPI: 1999-620161/53.

XX

XX N-PSDB; AAU31402.

XX

XX Novel polypeptides, used to develop products for the diagnosis,

PT prevention and treatment of pneumococcal infections

XX

XX Disclosure; Page 9; 85pp; English.

XX

XX The invention provides novel isolated polypeptides comprising the amino

CC acid sequence of an N-terminal choline binding protein A (Cbpa) truncate

CC (CbpaT). The polypeptides can be selected from sequences shown in

CC AAU9225, AAU9227, AAU9230, AAU9231, AAU9233 and AAU9238. The

CC polypeptides can be used for inducing an immune response in a subject

CC which has been exposed to or infected with a pneumococcal bacterium. They

CC can also be used for preventing infection by a pneumococcal bacterium.

CC Vaccines comprising the polypeptides or encoding nucleic acids can be

CC used for treating a subject infected with or exposed to a pneumococcal

CC bacterium. Antibodies specifically binding the polypeptides can be used

CC for detection and diagnosis and for preventing pneumococcal attachment to

CC a mucosal surface. The products can be used in humans and other animals

CC such as domestic animals, such as feline or canine subjects, farm animals

CC such as bovine, equine, caprine, ovine, and porcine subjects, wild

CC animals (whether in the wild or in a zoological garden), research

CC animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats,

CC i.e. for veterinary medical use.

XX

XX Sequence 655 AA:

QY

Best Local Similarity 1.5%; Score 10; DB 20; Length 655;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
 1111111111
 DB 113 ssssssdst 122

RESULT 26

AAV32099
 ID AAV32099 standard; Protein; 655 AA.

AC AAV32099;

DT 01-FEB-2000 (first entry)

DE Choline binding protein A (Cbpa).

KM Choline binding protein; Cbpa; adhesin; immunogen;

KW vaccine; diagnosis; therapy; pneumococcus; sepsis; meningitis;

OS Streptococcus pneumoniae.

PN WO951188-A2.

PD 14-OCT-1999.

PF 07-APR-1999; 99WO-US07669.

PR 07-APR-1998; 98US-0056019.

PR 07-APR-1998; 98US-0080878.

PA (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.

PI (MEDI-) MEDIMUNE INC.

DR WPI; 1999-633690/54.

DR N-PSDB; AA34452.

PT New N-terminal choline binding protein A truncate polypeptides, used to

PT develop products for the diagnosis, prevention and treatment of

PT pneumococcal infections

PS Disclosure; Page 88-93; 160pp; English.

CC This sequence represents the choline binding protein A (Cbpa)

CC of Streptococcus pneumoniae serotype type 4. The invention provides

CC novel N-terminal Cbpa truncated polypeptides (see AAV32098-110 and

CC AAV32179-90) and polynucleotides encoding them, host-vector systems,

CC and antibodies that specifically bind to the polypeptides. The

CC invention also relates to vaccines including the polypeptides, to

CC which provide protection or elicit protective antibodies to

CC bacterial infection, specifically pneumococcus, and to antibodies

CC and antagonists against such polypeptides for use in diagnosis and

CC passive immunotherapy. The polypeptides and/or polynucleotides are

CC also useful as competitive inhibitors of bacterial adhesin of

CC pneumococcus.

CC Sequence 655 AA;

Query Match 1.5%; Score 10; DB 20; Length 655;

Best Local Similarity 100.0%; Pred. No. 0.8;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
 1111111111

DB 113 ssssssdst 122

RESULT 27
 AAV81653
 ID AAV81653 standard; Protein; 694 AA.

XX AAV81653;
 AC
 XX
 DT 24-MAY-2000 (first entry)
 XX

DE Streptococcus pneumoniae protein sequence ID302.

KW Streptococcus pneumoniae; vaccine; screening; protein antigen;

KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;

KW pneumococcal disease.

OS Streptococcus pneumoniae.

PN WO200006737-A2.

PD 10-FEB-2000.

PF 27-JUL-1999; 99WO-GB02451.

PR 27-JUL-1998; 98GB-0016337.

PR 19-MAR-1999; 99US-0125164.

PA (MICR-) MICROBIAL TECHNICS LTD.

PI Gilbert CFG, Hansbro PM;

DR WPI; 2000-195300/17.

PT New Streptococcal protein, useful as a vaccine, for diagnosis of

PT pneumococcal diseases and for screening agents capable of antagonizing

PT or inhibiting expression of the protein

PS Claim 2; Page 96; 108pp; English.

CC AAV81501 to AAV81679 represent specifically claimed protein sequences

CC isolated from Streptococcus pneumoniae. AA05407 to AA05590 represent

CC specifically claimed nucleotide sequences isolated from S. pneumoniae.

CC The sequences have antibacterial and antiinflammatory properties.

CC The protein sequences, and fragments of them, are useful as immunogens

CC and/or antigens. The nucleotide sequences can be used in vaccines and in

CC diagnostic assays. The proteins and nucleotides can be useful for the

CC detection and diagnosis of S. pneumoniae. The protein sequences are also

CC useful for screening an agent capable of antagonizing, inhibiting or

CC interfering with the function or expression of the proteins in which the

CC agent is useful for treatment or prophylaxis of S. pneumoniae infection

CC and meningitis. AA05591 to AA05614 represent primers used in the

CC exemplification of the present invention.

CC Sequence 694 AA;

Query Match 1.5%; Score 10; DB 21; Length 694;

Best Local Similarity 100.0%; Pred. No. 0.84;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
 1111111111

DB 151 ssssssdst 160

RESULT 28

AAU36453
 ID AAU36453 standard; Protein; 774 AA.

AC AAU36453;

DT 14-FEB-2002 (first entry)

DE Pseudomonas aeruginosa cellular proliferation protein #443.

KW Antisense; prokaryotic cellular proliferation protein;

KW antibiotic; antibacterial; drug design.


```

OS   Pseudomonas aeruginosa.
PN   WO200170955-A2.
PD   27-SEP-2001.
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
PI   Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI   Yamamoto RT, Xu HH;
XX
XX   WPI; 2001-611495/70.
DR   N-PSDB; AAS54312.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
XX   Example 3; Seq ID No 12046; 511pp; English.
XX
XX   The invention relates to antisense inhibitors of genes essential to
XX   prokaryotic cellular proliferation, their use in identifying the
XX   genes, their use in the discovery of novel antibiotics, the essential
XX   themselves and the encoded proteins. The prokaryotes used are
XX   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX   invention is also useful for the identification of potential new targets
XX   for antibiotic development. The antisense nucleic acids can also be used
XX   to identify proteins used in proliferation, to express these proteins,
XX   and to obtain antibodies capable of binding to the expressed proteins.
XX   The proteins can be used to screen compounds in rational drug discovery
XX   programmes. The antisense nucleic acid sequence is also useful to screen
XX   for homologous nucleic acids which are required for cell proliferation in
XX   a wide variety of organisms. The present sequence represents an
XX   essential prokaryotic cellular proliferation protein.
XX
XX   Note: The sequence data for this patent did not form part
XX   of the printed specification, but was obtained in electronic
XX   format directly from WIPO at
XX   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence    774 AA;

Query Match          1.5%; Score 10; DB 22; Length 774;
Best Local Similarity 100.0%; Pred. NO. 0.93;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY      67 QGGSTLTGOL 76
        |||
Db       220 gggstltgql 229

RESULT  29
ID      AAU35684
XX      AAU35684 standard; Protein; 781 AA.
XX
XX      AAU35684;
DT      14-FEB-2002 (first entry)
DE      Haemophilus influenzae cellular proliferation protein #325.
XX
XX      Antisense; prokaryotic cellular proliferation protein;
XX      antibiotic; antibacterial; drug design.

```

```

xx Haemophilus influenzae.
os
xx WO200170955-A2.
xx
xx PD 27-SEP-2001.
xx
xx PE 21-MAR-2001; 2001WO-US09180.
xx
xx PR 21-MAR-2000; 2000US-191078P.
xx
xx PR 23-MAY-2000; 2000US-206848P.
xx
xx PR 26-MAY-2000; 2000US-207727P.
xx
xx PR 23-OCT-2000; 2000US-242578P.
xx
xx PR 27-NOV-2000; 2000US-253625P.
xx
xx PR 22-DEC-2000; 2000US-257931P.
xx
xx PR 16-FEB-2001; 2001US-269308P.
xx
xx PA (ELIT-) ELITRA PHARM INC.
xx
xx PL Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
xx Yamamoto RT, Xu HH;
xx WPI; 2001-611495/70.
xx DR N-PSDB; AAS53543.
xx
xx PT New polynucleotides for the identification and development of
xx antibiotics, comprise sequences of antisense nucleic acids -
xx
xx Example 3; Seq ID No 11277; 511pp; English.
xx
xx The invention relates to antisense inhibitors of genes essential to
xx prokaryotic cellular proliferation, their use in identifying the
xx genes, their use in the discovery of novel antibiotics, the essential
xx genes themselves and the encoded proteins. The prokaryotes used are
xx Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
xx pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
xx invention is also useful for the identification of potential new targets
xx for antibiotic development. The antisense nucleic acids can also be used
xx to identify proteins used in proliferation, to express these proteins,
xx and to obtain antibodies capable of binding to the expressed proteins.
xx The proteins can be used to screen compounds in rational drug discovery
xx programmes. The antisense nucleic acid sequence is also useful to screen
xx for homologous nucleic acids which are required for cell proliferation in
xx a wide variety of organisms. The present sequence represents an
xx essential prokaryotic cellular proliferation protein.
xx CC Note: The sequence data for this patent did not form part
xx of the printed specification, but was obtained in electronic
xx format directly from WIPO at
xx ftp.wipo.int/pub/published_pct_sequences.
xx
xx Sequence 781 AA;
xx
xx Query Match 1.5%; Score 10; DB 22; Length 781;
xx Best Local Similarity 100.0%; Pred. NO. 0.94;
xx Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
xx QY 67 OGGSTLTQOL 76
xx |||||||||
xx Db 218 gggstltqgl 227
xx
xx RESULT 30
xx ID AAR49657 standard; Protein; 1704 AA.
xx AC AAR49657;
xx DT 12-SEP-1994 (first entry)
xx DE Sequence of Heliothis armigera replicase encoded by RNA 1.
xx TW HASV; RNA 1; small RNA virus; replicase.

```


XX Heliothis armigera stunt virus.
 OS
 XX
 PN WO9404660-A.
 XX
 PD 03-MAR-1994.
 XX
 PF 13-AUG-1993; 93WO-AU00411.
 XX
 PR 14-AUG-1992; 92AU-0004081.
 PR 08-JUL-1993; 93US-0089372.
 XX
 PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
 XX (PACI-) PACIFIC SEEDS PTY LTD.
 PI Christian PD, Gordon KHJ, Hanzlik TN;
 XX
 DR WPI, 1994-083180/10.
 DR N-PSDB; AAQ58522.
 XX
 PT Small RNA virus capable of infecting insect species, e.g.
 PT Heliothis - and transgenic plants contg. viral nucleic acid, for
 PT protection against insect pests
 PS Disclosure: Figure 1, 183pp; English.
 XX
 XX The inventors claim a virus comprising a genome hybridisable with
 CC the nucleotide sequence of RNA 1 or RNA 2; pref. the sequences are
 CC those given in Figs 1 and 2 of the specification. Isolated proteins
 CC or polypeptide prepn. of the proteins or polypeptides derivable
 CC from the virus are also claimed.
 CC H. armigera larvae were raised and viral RNA was extracted. The virus
 CC RNAs were reverse transcribed into cDNA. Clone E3 represents 99.7%
 CC or RNA 1.(hr236 contains about 888 or RNA 2.)The full length clone
 CC of RNA 1 was completed using PCR. RNA 1 encodes a protein of mol.
 CC wt. 187,000 which is regarded as the replicase in view of its AA
 CC sequence similarity in certain limited regions to replicases of
 CC of other RNA viruses. The apparent mol. wt. of this protein upon
 CC in vitro translation of virus RNA and SDS-PAGE is 195,000. The
 CC sequence given in Figure 1 is in the same sense as the viral
 CC (positive-sense) RNA. There are other small ORfs at the 3' end,
 CC corresp. to the proteins P1a, P1b and P14.
 CC
 SQ Sequence 1704 AA;

Query Match 1.5%; Score 10; DB 15; Length 1704;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 617 SSSSDSSTS 626
 |||||||||
 DB 1452 ssssdsts 1461

Search completed: June 13, 2002, 08:46:22
 Job time: 263 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:45:24 ; Search time 21.51 Seconds
(without alignments)
2975.151 Million cell updates/sec

Title: US-08-961-083-2
Perfect score: 666
Sequence: 1 KTYDNKNQIADLGSERFVN.....TQGSNTTDPQGNQNPQAPQ 666

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 10

Total number of hits satisfying chosen parameters: 33

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	719	2 G95042	penicillin-binding
2	502	75.4	719	2 S28038	penicillin-binding
3	351	52.7	608	2 S28036	penicillin-binding
4	332	49.8	719	2 S28037	penicillin-binding
5	330	49.5	719	2 A42893	penicillin-binding
6	330	49.5	719	2 A97913	peptidoglycan glyc
7	181	27.2	719	2 S28031	penicillin-binding
8	181	27.2	719	2 S28034	penicillin-binding
9	181	27.2	719	2 S28033	penicillin-binding
10	181	27.2	719	2 S28032	penicillin-binding
11	97	14.6	188	2 S31952	penicillin-binding
12	79	11.9	719	2 S28035	penicillin-binding
13	53	8.0	637	2 B42893	penicillin-binding
14	20	3.0	139	2 S31941	penicillin-binding
15	14	2.1	664	2 G86692	penicillin-binding
16	12	1.8	760	2 E84953	penicillin-binding
17	11	1.7	801	2 B86673	penicillin-binding
18	11	1.7	821	2 E95245	penicillin-binding
19	11	1.7	821	2 B91110	peptidoglycan glyc
20	11	1.7	966	2 E84053	penicillin-binding
21	10	1.5	504	2 T29028	hypothetical prote
22	10	1.5	504	2 T21377	hypothetical prote
23	10	1.5	577	1 A36442	mitochondrial proc
24	10	1.5	693	2 H95255	choleline binding pr
25	10	1.5	730	2 A13460	penicillin-binding
26	10	1.5	757	2 G97472	penicillin-binding
27	10	1.5	757	2 AC2691	penicillin-binding
28	10	1.5	764	2 B97371	penicillin-binding
29	10	1.5	764	2 AB2589	penicillin-binding

30	10	1.5	774	2 H83057	penicillin-binding
31	10	1.5	777	2 A82303	penicillin-binding
32	10	1.5	781	2 D64138	penicillin-binding
33	10	1.5	824	2 AC0412	penicillin-binding

ALIGNMENTS

RESULT 1
G95042
penicillin-binding protein 1A [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C:Accession: G95042
R:Retcelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: G95042
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-719 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK74536.1; PID:914971838; GSPDB:GN00164; TIGR:
A:Experimental source: strain TIGR4
C:Gene: SP0369
C:Superfamily: penicillin-binding protein 1B

Query Match 100.0%; Score 666; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KTYDNKNQIADLGSERFVNQAQNDIPDVKATVSIEDHREFDHRGIDTITILGAFILN	60
DB	54	KTYDNKNQIADLGSERFVNQAQNDIPDVKATVSIEDHREFDHRGIDTITILGAFILN	113
QY	61	LOSNSLOGGSTLTFOOLIKLTFSTSTSDQTSIRKQEAWLAIQLDQKATKQELIYYINK	120
DB	114	LOSNSLOGGSTLTFOOLIKLTFSTSTSDQTSIRKQEAWLAIQLDQKATKQELIYYINK	173
QY	121	YVMSNGNTGMOTAQNYGKIDINLSLPQALLAAGPQAPNOYDPSHPEAAQDRRLIVL	180
DB	174	YVMSNGNTGMOTAQNYGKIDINLSLPQALLAAGPQAPNOYDPSHPEAAQDRRLIVL	233
QY	181	SEMKRQGYISAOYEKAVNPTITDGLQSLKSASNPAYMDNLTKEVINQVEBETGYNLT	240
DB	234	SEMKRQGYISAOYEKAVNPTITDGLQSLKSASNPAYMDNLTKEVINQVEBETGYNLT	293
QY	241	TGMADVTVNVDOQAOKHLMIDYNTDEVVAYPPDELOVASTYVYVNSGKYAQAQARHOSN	300
DB	294	TGMADVTVNVDOQAOKHLMIDYNTDEVVAYPPDELOVASTYVYVNSGKYAQAQARHOSN	353
QY	301	VSEGINQAVETNRDWSGTMKPTTDYAPALAEVGVYSTATIVHDEBYNPGTTPYVNMNR	360
DB	354	VSEGINQAVETNRDWSGTMKPTTDYAPALAEVGVYSTATIVHDEBYNPGTTPYVNMNR	413
QY	361	GYFGNITLQYALQOSRNPAYETLKKVGINRAKFTLNLGLIDYPSIHYSNALISSTTESD	420
DB	414	GYFGNITLQYALQOSRNPAYETLKKVGINRAKFTLNLGLIDYPSIHYSNALISSTTESD	473
QY	421	KKYGASSEKMAAAYAFANGGYTYPMYIHKVYPSDGSSEKESNNGTAMKETTYMMTD	480
DB	474	KKYGASSEKMAAAYAFANGGYTYPMYIHKVYPSDGSSEKESNNGTAMKETTYMMTD	533
QY	481	MKKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEELENIKTSQFVAPDELPAGYTRKYSM	540
DB	534	MKKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEELENIKTSQFVAPDELPAGYTRKYSM	593

OY 541 AWTGYSNRLTPLVNGLTVAAKVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFKNGAR 600
 |||||||
 Db 594 AWTGYSNRLTPLVNGLTVAAKVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFKNGAR 653
 |||||||
 OY 601 STWNSAPQOPSTESSSSSDSTSSSTSTSTNNSTTTNNNTQOQNTTPOQONON 660
 |||||||
 Db 654 STWNSAPQOPSTESSSSSDSTSSSTSTSTNNSTTTNNNTQOQNTTPOQONON 713
 |||||||
 OY 661 POPAP 666
 |||||||
 Db 714 POPAP 719

RESULT 2

S28038
 penicillin-binding protein 1a - Streptococcus pneumoniae (strain 45607) (fragment)
 C:Species: Streptococcus pneumoniae
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
 C:Accession: S28038
 R:Martin, C.; Sibold, C.; Hakenbeck, R.
 EMO J. 11, 3831-3836, 1992
 A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of pen
 A:Reference number: S28031; MIDID:93010977
 A:Status: translation not shown
 A:Accession: S28038
 A:Molecule type: DNA
 A:Residues: 1-719 <MAR>
 A:Cross-references: EMBL:X67873; NID:g47419; PIDN:CAA48073.1; PID:g47420
 C:Superfamily: penicillin-binding protein 1B

Query Match 75.4%; Score 502; DB 2; Length 719;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 602; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDITRILGAFARN 60
 |||||||
 Db 54 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDITRILGAFARN 113
 |||||||
 OY 61 LOSNSLQGGSTLTQOLIKLTFSTSDOTISKRAQEWALAIOLBKATKQELITYYINK 120
 |||||||
 Db 114 LOSNSLQGGSTLTQOLIKLTFSTSDOTISKRAQEWALAIOLBKATKQELITYYINK 173
 |||||||
 OY 121 VYMSNGNYGQTAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRLNLY 180
 |||||||
 Db 174 VYMSNGNYGQTAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRLNLY 233
 |||||||
 OY 181 SEKNNGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNYLKEYINOVEEETGYNLLT 240
 |||||||
 Db 234 SEKNNGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNYLKEYINOVEEETGYNLLT 293
 |||||||
 OY 241 TGMADVTTNVDQEAQKHMDIYNTDEVYVAYPPDELQVASTIVDSNGKVIAQLGARHOSN 300
 |||||||
 Db 294 TGMADVTTNVDQEAQKHMDIYNTDEVYVAYPPDELQVASTIVDSNGKVIAQLGARHOSN 353
 |||||||
 OY 301 VSFGINQAVETNRDWSGTMKPTIDYAPALEGYVDSTATIVHDEPNYPGTNTPVYNMDR 360
 |||||||
 Db 354 VSFGINQAVETNRDWSGTMKPTIDYAPALEGYVDSTATIVHDEPNYPGTNTPVYNMDR 413
 |||||||
 OY 361 GYFGNTTLOALQOQRNVAVERLNVGNLRATFLNGLIDYPSIHYSMAISSNTTESD 420
 |||||||
 Db 414 GYFGNTTLOALQOQRNVAVERLNVGNLRATFLNGLIDYPSIHYSMAISSNTTESD 473
 |||||||
 OY 421 KRYGASSEKMAAAYAFANGGTYYKPMYIHKVYSDGSEKEFSNVGTRAKKETATAMME 480
 |||||||
 Db 474 KRYGASSEKMAAAYAFANGGTYYKPMYIHKVYSDGSEKEFSNVGTRAKKETATAMME 533
 |||||||
 OY 481 MKKTVLTGTGRNAYLAMLPOAGKTGTSVYTDDEIENHIKTSOFVAPDELFAGYTRKYSM 540
 |||||||
 Db 534 MKKTVLTGTGRNAYLAMLPOAGKTGTSVYTDDEIENHIKTSOFVAPDELFAGYTRKYSM 593
 |||||||
 OY 541 AWTGYSNRLTPLVNGLTVAAKVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFKNGAR 600
 |||||||

Db 594 AWTGYSNRLTPLVNGLTVAAKVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFKNGAR 653
 |||||||
 OY 601 STW 603
 |||
 Db 654 STW 656

RESULT 3

S28036
 penicillin-binding protein 1a - Streptococcus pneumoniae (strain 8250) (fragment)
 C:Species: Streptococcus pneumoniae
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
 C:Accession: S28036
 R:Martin, C.; Sibold, C.; Hakenbeck, R.
 EMO J. 11, 3831-3836, 1992
 A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of
 A:Reference number: S28031; MIDID:93010977
 A:Accession: S28036
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-608 <MAR>
 A:Cross-references: EMBL:X67871; NID:g47415; PIDN:CAA48071.1; PID:g47416
 C:Superfamily: penicillin-binding protein 1B

Query Match 52.7%; Score 351; DB 2; Length 608;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDITRILGAFARN 60
 |||||||
 Db 54 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDITRILGAFARN 113
 |||||||
 OY 61 LOSNSLQGGSTLTQOLIKLTFSTSDOTISKRAQEWALAIOLBKATKQELITYYINK 120
 |||||||
 Db 114 LOSNSLQGGSTLTQOLIKLTFSTSDOTISKRAQEWALAIOLBKATKQELITYYINK 173
 |||||||
 OY 121 VYMSNGNYGQTAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRLNLY 180
 |||||||
 Db 174 VYMSNGNYGQTAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRLNLY 233
 |||||||
 OY 181 SEKNNGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNYLKEYINOVEEETGYNLLT 240
 |||||||
 Db 234 SEKNNGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNYLKEYINOVEEETGYNLLT 293
 |||||||
 OY 241 TGMADVTTNVDQEAQKHMDIYNTDEVYVAYPPDELQVASTIVDSNGKVIAQLGARHOSN 300
 |||||||
 Db 294 TGMADVTTNVDQEAQKHMDIYNTDEVYVAYPPDELQVASTIVDSNGKVIAQLGARHOSN 353
 |||||||
 OY 301 VSFGINQAVETNRDWSGTMKPTIDYAPALEGYVDSTATIVHDEPNYPGT 351
 |||||||
 Db 354 VSFGINQAVETNRDWSGTMKPTIDYAPALEGYVDSTATIVHDEPNYPGT 404
 |||||||

RESULT 4

S28037
 penicillin-binding protein 1a - Streptococcus pneumoniae (strain 63915) (fragment)
 C:Species: Streptococcus pneumoniae
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
 C:Accession: S28037
 R:Martin, C.; Sibold, C.; Hakenbeck, R.
 EMO J. 11, 3831-3836, 1992
 A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of
 A:Reference number: S28031; MIDID:93010977
 A:Accession: S28037
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-719 <MAR>
 A:Cross-references: EMBL:X67872; NID:g47417; PIDN:CAA48072.1; PID:g47418
 C:Superfamily: penicillin-binding protein 1B

Query Match 49.8%; Score 332; DB 2; Length 719;

Best Local Similarity 100.0%; Pred. No. 0; Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNOLLIDLCSERRVNAQANDIPDLVKAIVSIEDHFFDRGIDITRILGAFLRN 60
DB 54 KIYDNKNOLLIDLCSERRVNAQANDIPDLVKAIVSIEDHFFDRGIDITRILGAFLRN 113
QY 61 LOSNLSAGSGFLTQOOLIKLYFFSTSDOTSIRKAOEAMLAIOLEOKATKOEILTYINK 120
DB 114 LOSNLSAGSGFLTQOOLIKLYFFSTSDOTSIRKAOEAMLAIOLEOKATKOEILTYINK 173
QY 121 VYMSNGNGMTAAQNTYNGKDLNNISLPQALLAGMPAPNOYDPSHPEAAODRRNLVL 180
DB 174 VYMSNGNGMTAAQNTYNGKDLNNISLPQALLAGMPAPNOYDPSHPEAAODRRNLVL 233
QY 181 SEMKNOGYTSAEYKAVNTPTDGLQSLKASNPAYMDNLYLKEVINQVEETGYNLLT 240
DB 234 SEMKNOGYTSAEYKAVNTPTDGLQSLKASNPAYMDNLYLKEVINQVEETGYNLLT 293
QY 241 TGMVYTNVDOEAQKHLMDIYNTDEYVAPDDELQVASTIVDSNGKYIAOLGARHOSN 300
DB 294 TGMVYTNVDOEAQKHLMDIYNTDEYVAPDDELQVASTIVDSNGKYIAOLGARHOSN 353
QY 301 VSPGINQAVETNRDWSGTMRPTDYAPALEYG 332
DB 354 VSPGINQAVETNRDWSGTMRPTDYAPALEYG 385
RESULT 5
A42893
penicillin-binding protein 1A - Streptococcus pneumoniae
C:Species: Streptococcus pneumoniae
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 21-Jan-2000
C:Accession: A42893
R:Martin, C.; Briese, T.; Hakenbeck, R.
J. Bacteriol. 174, 4517-4523, 1992
A:Title: Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus pneumoniae
A:Reference number: A42893; MUID:92325042
A:Accession: A42893
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: GB:M90527; NID:g153766; PIDN:AAA26956.1; PID:g153768
C:Superfamily: penicillin-binding protein 1B

Query Match 49.5%; Score 330; DB 2; Length 719;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 530; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 72 LTQOLIKLYTSTSTSDOTSIRKAOEAMLAIOLEOKATKOEILTYINKVYMSNGNYGMQ 131
DB 125 LTQOLIKLYTSTSTSDOTSIRKAOEAMLAIOLEOKATKOEILTYINKVYMSNGNYGMQ 184
QY 132 TAAQNTYNGKDLNNISLPQALLAGMPAPNOYDPSHPEAAODRRNLVLSEKKNQYISA 191
DB 185 TAAQNTYNGKDLNNISLPQALLAGMPAPNOYDPSHPEAAODRRNLVLSEKKNQYISA 244
QY 192 EYKAVNTPTDGLQSLKASNPAYMDNLYLKEVINQVEETGYNLLTGMVYTNVDO 251
DB 245 EYKAVNTPTDGLQSLKASNPAYMDNLYLKEVINQVEETGYNLLTGMVYTNVDO 304
QY 252 EAKHLMIDYNTDEYVAPDDELQVASTIVDSNGKYIAOLGARHOSNVSFGINQAVET 311
DB 305 EAKHLMIDYNTDEYVAPDDELQVASTIVDSNGKYIAOLGARHOSNVSFGINQAVET 364
QY 312 NRDMGSTMKPTTDYAPALEYGVYDSTATIVHDEPNYNGTNPYVNMGRGYFGNITLQYA 371
DB 365 NRDMGSTMKPTTDYAPALEYGVYDSTATIVHDEPNYNGTNPYVNMGRGYFGNITLQYA 424
QY 372 LQOSRNVPVAVETLKNKVGINRAKTFPLNGCIDYPSIHYSNAISSNTESDKKYGASSEKMA 431
DB 431 LQOSRNVPVAVETLKNKVGINRAKTFPLNGCIDYPSIHYSNAISSNTESDKKYGASSEKMA 484

DB 425 LQOSRNVPVAVETLKNKVGINRAKTFPLNGCIDYPSIHYSNAISSNTESDKKYGASSEKMA 484
QY 432 AAYAFANGGTYRPMYTHKVVSDGSEKESNVGTRAMKETTAAYMTDMKKTIVLTGTG 491
DB 485 AAYAFANGGTYRPMYTHKVVSDGSEKESNVGTRAMKETTAAYMTDMKKTIVLTGTG 544
QY 492 RNAYLAWLPQAGKTGTSNVTDEIEHNHKTQSOFVAPDELFAGYTRKYSMAVWTGYSNRLT 551
DB 545 RNAYLAWLPQAGKTGTSNVTDEIEHNHKTQSOFVAPDELFAGYTRKYSMAVWTGYSNRLT 604
QY 552 PLVNGTLTVAAKVYRSMNTYLSGSPEDWNPBELYNGSERVFKNGARSTW 603
DB 605 PLVNGTLTVAAKVYRSMNTYLSGSPEDWNPBELYNGSERVFKNGARSTW 656

RESULT 6
A97913
peptidoglycan glycosyltransferase (EC 2.4.1.129) [imported] - Streptococcus pneumonia
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: A97913
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAnthen, S.;
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A97913
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-719 <KUR>
A:Cross-references: GB:A607317; PIDN:AAK9133.1; PID:g15457886; GSPDB:GN00174
C:Genetics:
A:Gene: pbpA
C:Superfamily: penicillin-binding protein 1B
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 49.5%; Score 330; DB 2; Length 719;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 530; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 72 LTQOLIKLYTSTSTSDOTSIRKAOEAMLAIOLEOKATKOEILTYINKVYMSNGNYGMQ 131
DB 125 LTQOLIKLYTSTSTSDOTSIRKAOEAMLAIOLEOKATKOEILTYINKVYMSNGNYGMQ 184
QY 132 TAAQNTYNGKDLNNISLPQALLAGMPAPNOYDPSHPEAAODRRNLVLSEKKNQYISA 191
DB 185 TAAQNTYNGKDLNNISLPQALLAGMPAPNOYDPSHPEAAODRRNLVLSEKKNQYISA 244
QY 192 EYKAVNTPTDGLQSLKASNPAYMDNLYLKEVINQVEETGYNLLTGMVYTNVDO 251
DB 245 EYKAVNTPTDGLQSLKASNPAYMDNLYLKEVINQVEETGYNLLTGMVYTNVDO 304
QY 252 EAKHLMIDYNTDEYVAPDDELQVASTIVDSNGKYIAOLGARHOSNVSFGINQAVET 311
DB 305 EAKHLMIDYNTDEYVAPDDELQVASTIVDSNGKYIAOLGARHOSNVSFGINQAVET 364
QY 312 NRDMGSTMKPTTDYAPALEYGVYDSTATIVHDEPNYNGTNPYVNMGRGYFGNITLQYA 371
DB 365 NRDMGSTMKPTTDYAPALEYGVYDSTATIVHDEPNYNGTNPYVNMGRGYFGNITLQYA 424
QY 372 LQOSRNVPVAVETLKNKVGINRAKTFPLNGCIDYPSIHYSNAISSNTESDKKYGASSEKMA 431
DB 425 LQOSRNVPVAVETLKNKVGINRAKTFPLNGCIDYPSIHYSNAISSNTESDKKYGASSEKMA 484
QY 432 AAYAFANGGTYRPMYTHKVVSDGSEKESNVGTRAMKETTAAYMTDMKKTIVLTGTG 491
DB 485 AAYAFANGGTYRPMYTHKVVSDGSEKESNVGTRAMKETTAAYMTDMKKTIVLTGTG 544
QY 492 RNAYLAWLPQAGKTGTSNVTDEIEHNHKTQSOFVAPDELFAGYTRKYSMAVWTGYSNRLT 551
DB 545 RNAYLAWLPQAGKTGTSNVTDEIEHNHKTQSOFVAPDELFAGYTRKYSMAVWTGYSNRLT 604

Db 545 RNATLAWLPQAGKGTGSNYDDEELENNIKTSQFAPEDELPAAGTYRKKSMAWVIGYSNRLLT 604

Qy 552 PLVNGGLTVAAYKYYRSMYTYLSEGSNPEDNNIPGGLYRNGEYFVKNGARSTW 603
|||||
605 PLVNGGLTVAAYKYYRSMYTYLSEGSNPEDNNIPGGLYRNGEYFVKNGARSTW 656

penicillin-binding protein 1a - *Streptococcus pneumoniae* (strain 456) (fragment)

C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: 528031
R:Martin, C.; Stibold, C.; Hakenbeck, R.
EMBO J. 11,3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant *Staphylococcus aureus* strains
A:Reference number: 528031; MUID:93010977
A:Accession: 528031
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67868; NID:947405; PIDN:CAA8068.1; PID:947406
A:Superfamily: penicillin-binding protein 1B

Query Match	27.2%;	Score 181;	DB 2;	Length 719;
Best Local Similarity	100.0%;	Pred. No. 5.5e-175;		
Matches 181; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	51	IRIGAEFLRNLOQSNLSLOGGSLTFOQLIKLTYFSTSDQYISRAQDAWALDLEOKATK	110
Db	104	IRIGAEFLRNLOQSNLSLOGGSLTFOQLIKLTYFSTSDQYISRAQDAWALDLEOKATK	163
QY	111	QELITYYINKVYMSNGNYGMQTAQNYTGYKDLNNLSLPQALLAGMFAQBNQYDPYSHE	170
Db	164	QELITYYINKVYMSNGNYGMQTAQNYTGYKDLNNLSLPQALLAGMFAQBNQYDPYSHE	223
QY	171	AAQDRNLVLESMKNQGYISAEQYKAVNPIIDGLOSLKSASNPAYMYMNTYKQYINQY	230
Db	224	AAQDRNLVLESMKNQGYISAEQYKAVNPIIDGLOSLKSASNPAYMYMNTYKQYINQY	283
QY	231	E 231	
Db	284	E 284	

RESULT 8
S28034
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 56742) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28034
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant strains of *Streptococcus pneumoniae*
A:Reference number: S28031; MIMD:93010977
A:Accession: S28034
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67869; NID:q47411; PIDN:CAAG8065.1; PID:q47412
A:Superfamily: penicillin-binding protein 1B

Query Match	27.2%;	Score 181;	DB 2;	Length 719;
Best Local Similarity	100.0%;	Pred. No. 5.5e-175;		
Matches 181;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

Db	164	QELITYINKYMSNGVNTAAQNTYGGKDLNLSLPQLALLAGMGAQANQDPRYSHPRE	2233
Qy	171	AAODRRNLVLSSEMNQOYISAEOYEKAVNTPITDGLQSLKSASNYPAYMDNYLKEVINQV	230
Db	224	AAODRRNLVLSSEMNQOYISAEOYEKAVNTPITDGLQSLKSASNYPAYMDNYLKEVINQV	283

QY	231	E	231
		-	
Db	284	E	284

RESULT 9
S28033
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 670) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28033
R:Martin, C., Sibold, C., Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of
A:Reference number: S28031; MUID:93010977
A:Accession: S28033
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67867; NID:g47409; PIDN:CAA48067.1; PID:g47410
C:Superfamily: penicillin-binding protein 1B

Query Match	27.28;	Score 181;	DB 2;	Length 719;
Best Local Similarity	100.0%;	Pred. No. 5.5e-175;		
Matches 181;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	51	IRIGAEFLRNLOSMLOGGSLTFQOLIKLTFYSTSDQYISKRABEAWIALOLEKATK	110
Db	104	IRIGAEFLRNLOSMLOGGSLTFQOLIKLTFYSTSDQYISKRABEAWIALOLEKATK	163
Qy	111	QELITFYINKVYMSNGYMGOTAAQNTYGGKDLNNLSLPOLALLAGMPOABNOYDYSHP	170
Db	164	QELITFYINKVYMSNGYMGOTAAQNTYGGKDLNNLSLPOLALLAGMPOABNOYDYSHP	223
Qy	171	AAQORRNVLSEMNQGYISAEQYERKAVNPIPIIDGLOSLKSASNYPAYMONTYLAKEYINQY	230
Db	224	AAQORRNVLSEMNQGYISAEQYERKAVNPIPIIDGLOSLKSASNYPAYMONTYLAKEYINQY	283
Qy	231	E 231	
Db	284	E 284	

RESULT 10
S28032
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 681) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28032
R:Martin, C., Sibold, C., Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of
A:Reference number: S28031; MUID:93010977
A:Accession: S28032
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67866; NID:g47407; PIDN:CAA48066.1; PID:g47408
A:Superfamily: penicillin-binding protein 1B

Query Match	27.2%	Score 181;	DB 2;	Length 719;
Best Local Similarity	100.0%	Pred. No. 5.5e-175;		
Matches 181; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	51	IRIIAFLRNLOSLSLGGSTLTLOOLIKLYTSTSTSTOOTISRKQOEMALATOLEOKATK	110
Db	104	IRILFAFLRNLOSLSLGGSTLTLOOLIKLYTSTSTSTOOTISRKQOEMALATOLEOKATK	163
QY	111	QELIYYIINKVYMSGNGYMGTAQNTYYGKDLNNLSLPOLALLAGMPOAPNOYDEYSHE	170
Db	164	QELIYYIINKVYMSGNGYMGTAQNTYYGKDLNNLSLPOLALLAGMPOAPNOYDEYSHE	223
QY	171	AAQDRRNLTVLESMKNQGYISAEQYERAVNTPTDGLQSLKSASNPAYMDNLYKEVINQY	230
Db	224	AAQDRRNLTVLESMKNQGYISAEQYERAVNTPTDGLQSLKSASNPAYMDNLYKEVINQY	283
QY	231	E 231	
Db	284	E 284	

RESULT 11
S31952
penicillin-binding protein 1A - Streptococcus pneumoniae (strain 85983) (fragment)
C:Species: Streptococcus pneumoniae
A:Variety: strain 85983
C:Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 26-May-2000
C:Accession: S31952
R:Kell, C.M.; Jordens, Z.; Daniels, M.; Coffey, T.J.; Bates, J.; Paul, J.; Gilks, C.; Segal, J.
submitted to the EMBL data library, February 1993
A:Description: Molecular epidemiology of penicillin-resistant pneumococci isolated in New Zealand
A:Reference number: S31941

A: Molecule type: DNA
A: Residues: 1-188 <KE>
A: Cross-references: EMBL: Z21800
A: Experimental source: strain 85983
C: Superfamily: penicillin-binding protein 1B
C: Keywords: antibiotic resistance; penicillin resistance

Query Match	14.6%	Score 97	DB 2	Length 188
Best Local Similarity	100.0%	Pred. No.	3.4e-90	
Matches 97	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0

Oy	557	GLTAAKAYRRMMNTLYLSCGSPEDMNIPEGILYRNGEVEFKNGARSTWNSAPQOPSTES	61.6
Db	92	GLTAAKAYRRMMNTLYLSCGSPEDMNIPEGILYRNGEVEFKNGARSTWNSAPQOPSTES	151.1
Oy	617	SSSSSDSTSOSSSTPSTNNSTTTNPNNNTQOOSTT	653
Db	152	SSSSSDSTSOSSSTPSTNNSTTTNPNNNTQOOSTT	188

RESULT 12
S28035
penicillin-binding protein 1A - Streptococcus pneumoniae (strain 2039) (fragment)
C:Species: Streptococcus pneumoniae
A:Variety: strain 2039
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28035
R:Martin, C., Sibbold, C., Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant strains of Streptococcus pneumoniae
A:Reference number: S28031; MUID:93010977
A:Accession: S28035
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67870; NID:g47413; PIDN:CAAA8070.1; PID:g47414
C:Experimental source: strain 2039
A:Superfamily: penicillin-binding protein 1B
;Keywords: antibiotic resistance; penicillin resistance

Query Match	Score	DB	Length
11.98;	79;	2;	719;
Best Local Similarity	Pred. No.	2.2e-71;	

Matches	79: Conservative	0: Mismatches	0: Indels	0: Gaps
QY 64	NSLOGSGSTLQOLIKITVSTSNQITSRKQEAALAIQLEOKAKRQELIYYINKVM			123
Db 117	NSLOGSGSTLQOLIKITVSTSNQITSRKQEAALAIQLEOKAKRQELIYYINKVM			176
QY 124	SNQNTGMOTAAQNYGKDL	142		
Db 177	SNQNTGMOTAAQNYGKDL	195		

RESULT 13
BA2893
penicillin-binding protein 1 - Streptococcus orlilis
C:Species: Streptococcus orlilis
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 04-Mar-2000
C:Accession: BA2893
R:Martin, C.; Briese, T.; Hakenbeck, R.
J. Bacteriol. 174, 4517-4523, 1992
A:Title: Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus orlilis
A and 1B.

A;Residues: 1-637 <MAR>
A;Cross-references: GB:M90528; NID:q153769; PIDN:AAA26958.1; PID:q1537711
C;Superfamily: penicillin-binding protein 1B

Query Match	8.0%	Score 53	DB 2	Length 637
Best Local Similarity	100.0%	Pred. No.	4.9e-45	
Matches 53	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY 286 GKVAQLGARHSSNVSEFGINAVETNRDMGSTMKPTIDYAPALECYVDSTA 338
 |||||
 Db 340 GKVAQLGARHSSNVSEFGINAVETNRDMGSTMKPTIDYAPALECYVDSTA 392

RESULT 14
S31941
penicillin-binding protein 1A - Streptococcus pneumoniae (strain 100511) (fragment)
C:Species: Streptococcus pneumoniae
A:Variety: strain 100511
C:Date: 06-Jan-1995 #sequence-revision 06-Jan-1995 #text-change 26-May-2000
C:Accession: S31941
R:Kell, C.M.; Jordens, Z.; Daniels, M.; Coffey, T.J.; Bates, J.; Paul, J.; Gilks, C.;
submitted to the EMBL data library, February 1993
A:Description: Molecular epidemiology of penicillin-resistant pneumococci isolated in
A:Reference number: S31941
A:Accession: S31941
A:Molecule type: DNA
A:Residues: 1-139 <KE>
A:Cross-references: EMBL:Z21809; NID:949384; PIDN:CAA79874.1; PID:9940871
A:Experimental source: strain 100511
C:Superfamily: penicillin-binding protein 1B
C:Keywords: antibiotic resistance; penicillin resistance

Query Match	3.0%	Score 20;	DB 2;	Length 139;
Best Local Similarity	100.0%	Pred. No. 4.1e-12;		
Matches 20; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 533 GYTRKYSMAVWTGYSNRLTP 552
 |||||
 Db 16 GYTRKYSMAVWTGYSNRLTP 35

```

RESULT      15
GB6692
penicillin-binding protein 1A [imported] - Lactococcus lactis subsp. lactis (strain I
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

```


C:Accession: G86692
R:Botlittin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: G86692
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-664 <STO>
A:Cross-references: GB:AE005176; PID:g12723428; PIDN:AAK04641.1; GSPDB:GN00146
A:Experimental source: strain ILL403
C:Genetics:
A:Gene: pona
C:Superfamily: penicillin-binding protein 1B

Query Match 2.1%; Score 14; DB 2; Length 664;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 GASSEKMAAYAF 437
|||||
DB 467 GASSEKMAAYAF 480

RESULT 16
E84953
penicillin-binding protein 1b [imported] - *Buchnera* sp. (strain APS)
C:Species: *Buchnera* sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: E84953
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. A
A:Reference number: A84930; MUID:20445173
A:Accession: E84953
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-760 <STO>
A:Cross-references: GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: mrcB; BU200

Query Match 1.8%; Score 12; DB 2; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 QGGSTLTQOLIK 78
|||||
DB 204 QGGSTLTQOLIK 215

RESULT 17
B86673
penicillin-binding protein 1b [imported] - *Lactococcus lactis* subsp. *lactis* (strain ILL4
C:Species: *Lactococcus lactis* subsp. *lactis*
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: B86673
R:Botlittin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: B86673
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-801 <STO>
A:Cross-references: GB:AE005176; PID:g12723258; PIDN:AAK04484.1; GSPDB:GN00146
A:Experimental source: strain ILL403
C:Genetics:
A:Gene: pbp1b

Query Match 1.7%; Score 11; DB 2; Length 801;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSTLTQOLIK 78
|||||
DB 172 GGSTLTQOLIK 182

RESULT 18
E95245
penicillin-binding protein 1b [imported] - *Streptococcus pneumoniae* (strain TIGR4)
C:Species: *Streptococcus pneumoniae*
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: E95245
R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapp
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A:Title: Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*.
A:Reference number: A95000; MUID:21357209; PMID:11453916
A:Accession: E95245
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-821 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK76158.1; PID:g14973609; GSPDB:GN00164; TIGR
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP2099

Query Match 1.7%; Score 11; DB 2; Length 821;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSTLTQOLIK 78
|||||
DB 176 GGSTLTQOLIK 186

RESULT 19
B99110
peptidoglycan glycosyltransferase (EC 2.4.1.129) [imported] - *Streptococcus pneumonia*
C:Species: *Streptococcus pneumoniae*
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: B99110
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; Mcahren, S.
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaaskunas, S.
A:Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: B99110
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-821 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAL00711.1; PID:g15459604; GSPDB:GN00174
A:Genetics:
A:Gene: pbp1b
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 1.7%; Score 11; DB 2; Length 821;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSTLTQOLIK 78
|||||
DB 176 GGSTLTQOLIK 186

RESULT 20
E84053
penicillin-binding protein 1A pdpf [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: E84053
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: E84053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-966 <STO>
A:Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BA06948.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: pdpf

Query Match 1.7%; Score 11; DB 2; Length 966;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSTLTQOLIK 78
|||||
DB 157 GGSTLTQOLIK 167

RESULT 21
T29028
hypothetical protein F53612.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29028
R:Wu, X.; Graves, T.
Submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid F53612.
A:Reference number: Z20555
A:Accession: T29028
A:Cross-references: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-294 <WUX>
A:Status: preliminary
A:Experimental source: strain Bristol N2; clone F53612
C:Genetics:
A:Gene: CESP:F53612.4
A:Map position: 1
A:Introns: 111/3; 271/1

Query Match 1.5%; Score 10; DB 2; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 617 SSSSSSDSTS 626
|||||
DB 126 SSSSSSDSTS 135

RESULT 22
T21377
hypothetical protein F25H9.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T21377
R:Smey, R.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z19414
A:Accession: T21377
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-504 <ML>
A:Cross-references: EMBL:Z81069; PIDN:CAB02995.1; GSPDB:GN00023; CESP:F25H9.5
A:Experimental source: clone F25H9
C:Genetics:
A:Gene: CESP:F25H9.5
A:Map position: 5
A:Introns: 7/1; 34/3; 88/1; 189/3; 277/2; 343/3; 442/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F25H9.5

Query Match 1.5%; Score 10; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 617 SSSSSSDSTS 626
|||||
DB 443 SSSSSSDSTS 452

RESULT 23
A36442
mitochondrial processing peptidase (EC 3.4.24.64) alpha chain precursor - Neurospora
N:Alternate names: alpha-MPP; mitochondrial processing peptidase catalytic chain; MPP
C:Species: Neurospora crassa
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 19-Jan-2001
C:Accession: A36442; S36362
R:Schneider, H.; Arretz, M.; Machter, E.; Neupert, W.
J. Biol. Chem. 265, 9881-9887, 1990
A:Title: Matrix processing peptidase of mitochondria. Structure-function relationship
A:Reference number: A36442; MUID:90277682
A:Accession: A36442
A:Molecule type: mRNA
A:Residues: 1-577 <SCH1>
A:Cross-references: GB:J05484
R:Schneider, H.; Arretz, M.; Machter, E.; Neupert, W.
Submitted to the EMBL Data Library, July 1990
A:Reference number: S36362
A:Accession: S36362
A:Molecule type: mRNA
A:Residues: 1-106, A, 108-577 <SCH2>
A:Cross-references: GB:J05484; NID:g168840; PIDN:AAA33597.1; PID:g168841
C:Genetics:
A:Gene: MPP
A:Genome: nuclear
C:Superfamily: mitochondrial processing peptidase alpha chain
C:Keywords: heterodimer; hydrolase; metalloproteinase; mitochondrial matrix; mitochon
F:1-35/Domain: transit peptide (mitochondrion) #status predicted <TP>
F:36-577/Product: mitochondrial processing peptidase alpha chain #status predicted <M

Query Match 1.5%; Score 10; DB 1; Length 577;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 615 ESSSSSDSS 624
|||||
DB 286 ESSSSSDSS 295

RESULT 24
H95255
choline binding protein A [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: H95255
R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
son, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf
son, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: H95255

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-693 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK76241.1; PID:g14973701; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP2190

Query Match 1.5%; Score 10; DB 2; Length 693;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
|||||
DB 151 SSSSSSDSST 160

RESULT 25
A13480
penicillin-binding protein 1A [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: A13480
R:DeVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, R.; Mazur, M.; Goldstein, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AB2252; PMID:11756688
A:Accession: A13480
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-730 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL5012.1; PID:g17963868; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME1831
A:Map position: 1

Query Match 1.5%; Score 10; DB 2; Length 730;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 QGGSTLTQOL 76
|||||
DB 146 QGGSTLTQOL 155

RESULT 26
G97472
penicillin-binding protein 1A [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: G97472
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; PMID:11743194
A:Accession: G97472
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-757 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK6736.1; PID:g15155930; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_1696
A:Map position: circular chromosome

Query Match 1.5%; Score 10; DB 2; Length 757;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 QGGSTLTQOL 76
|||||
DB 193 QGGSTLTQOL 202

RESULT 27
AC2691
penicillin binding protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AC2691
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl ; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AC2691
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-757 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL41945.1; PID:g17739313; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu0931
A:Map position: circular chromosome

Query Match 1.5%; Score 10; DB 2; Length 757;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 QGGSTLTQOL 76
|||||
DB 193 QGGSTLTQOL 202

RESULT 28
B97371
penicillin-binding protein 1A [imported] - Agrobacterium tumefaciens (strain C58, Cer)
C:Species: Agrobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: B97371
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; PMID:11743194
A:Accession: B97371
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-764 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK85923.1; PID:g15154974; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_158
A:Map position: circular chromosome

Query Match 1.5%; Score 10; DB 2; Length 764;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 QGGSTLTQOL 76
|||||
DB 183 QGGSTLTQOL 192

RESULT 29
AB2589
penicillin-binding protein pbpc [imported] - Agrobacterium tumefaciens (strain C58, D

C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AB2589
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCell
; Kap, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AB2589
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-764 <KUR>
A:Cross-references: GB:AE008688; PIDN:AL41128.1; PID:g17738422; GSPDB:GN00186
C:Genetics:
A:Experimental source: strain C58 (Dupont)
A:Gene: pbpc
A:Map position: circular chromosome

Query Match 1.5%; Score 10; DB 2; Length 764;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 OGGSTLTOOL 76
|||||
Db 183 OGGSTLTOOL 192

RESULT 30
H83057
penicillin-binding protein 1B PA4700 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83057
R:Stover, C.K.; Plam, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337
A:Accession: H83057
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-774 <STO>
A:Cross-references: GB:AE004884; GB:AE004091; NID:g9950954; PIDN:AG08086.1; GSPDB:GN001
C:Genetics:
A:Experimental source: strain PA01
A:Gene: mrcB; PA4700

Query Match 1.5%; Score 10; DB 2; Length 774;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 OGGSTLTOOL 76
|||||
Db 220 OGGSTLTOOL 229

RESULT 31
AB2303
penicillin-binding protein 1B VC0602 [imported] - Vibrio cholerae (strain N16961 serogro
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: AB2303
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, R
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: AB2035; MUID:20406833
A:Accession: AB2303
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-777 <HEI>
A:Cross-references: GB:AE004145; GB:AE003852; NID:g9655032; PIDN:AAF93769.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0602
A:Map position: 1
C:Superfamily: penicillin-binding protein 1B

Query Match 1.5%; Score 10; DB 2; Length 777;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 OGGSTLTOOL 76
|||||
Db 230 OGGSTLTOOL 239

RESULT 32
D64138
penicillin-binding protein 1B homolog - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 21-Jul-2000
C:Accession: D64138
R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Feldman
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A6400; MUID:9530630
A:Accession: D64138
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-781 <TIG>
A:Cross-references: GB:U02845; GB:142023; NID:g3212236; PIDN:AAC23371.1; PID:g1574581
C:Superfamily: penicillin-binding protein 1B

Query Match 1.5%; Score 10; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 OGGSTLTOOL 76
|||||
Db 218 OGGSTLTOOL 227

RESULT 33
AC0412
penicillin-binding protein 1B (EC 2.4.2.-) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C:Accession: AC0412
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Ritball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barril
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AC0412
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-824 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92623.1; PID:g15981319; GSPDB:GN00175
C:Genetics:
A:Gene: mrcB
C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 1.5%; Score 10; DB 2; Length 824;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 67 QGGSTLTQOL 76
 |||||
 Db 266 QGGSTLTQOL 275

Search completed: June 13, 2002, 08:49:14
 Job time: 230 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:46:45 ; Search time 13.57 seconds
(without alignments)
1900.311 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 666
Sequence: 1 KIYDNKQIADLGSERYN.....TQSSNTTPQOQNQPPAP 666

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 10

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : SwissProt.40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	719	1	PPPA_STRPN
2	53	8.0	637	1	PPPA_STROR
3	12	1.8	760	1	PPPB_BUCAI
4	10	1.5	577	1	MPPA_NEUCR
5	10	1.5	777	1	PPPB_VIBCH
6	10	1.5	781	1	PPPB_HAEIN

ALIGNMENTS

RESULT 1
ID PPBA_STRPN STANDARD; PRT; 719 AA.
AC Q04707;
DT 01-OCT-1994 (Rel. 30, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Penicillin-binding protein 1A (PBP-1A) (Exported protein 2).
GN PONA OR EXP2 OR SP0369.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN=45607, AND 63915;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain.";
RL EMBL J. 11:3831-3836(1992).

[2]
RP SEQUENCE FROM N.A.
RC STRAIN-R6;
RX MEDLINE=92325042; PubMed=1624444;
RA Martin C., Briese T., Hakenbeck R.;
RT "Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus pneumoniae and Streptococcus oralis with high homology to Escherichia coli penicillin-binding proteins 1a and 1b.";
RL J. Bacteriol. 174:4517-4523(1992).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN-TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D., Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J., Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D., Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D., Holtzapple E., Khouli H., Wolf A.M., Uitterback T.R., Hansen C.L., McDonald L.A., Feldblyum T.V., Angiolini S., Dickinson T., Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C., Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.";
RL Science 293:498-506(2001).
[4]
RP SEQUENCE OF 293-369 FROM N.A.
RC STRAIN-R6X;
RX MEDLINE=95020625; PubMed=7934910;
RA Pearce B.J., Yin Y.B., Masure H.R.;
RT "Genetic identification of exported proteins in Streptococcus pneumoniae.";
RL Mol. Microbiol. 9:1037-1050(1993).
CC -I- FUNCTION: CELL WALL FORMATION.
CC -I- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- MISCELLANEOUS: THESE STRAINS ARE PENICILLIN-SENSITIVE.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X67873; CAA48073.1; -;
CC EMBL: X67872; CAA48072.1; -;
CC EMBL: M90527; AAA26956.1; -;
CC EMBL: AE007349; AAK74536.1; -;
CC PIR: S28038; S28038.
CC TIGR: SP0369; -;
CC InterPro: IPR001264; Transglycosyl.
CC InterPro: IPR001460; Transpeptidase.
CC Pfam: PF00912; Transglycosyl. 1.
CC Pfam: PF00905; Transpeptidase; 1.
CC ProDom: PD001895; Transglycosyl. 1.
CC Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
CC Multifunctional enzyme; Complete proteome.
CC ACCT_SITE 370 370
CC DOMAIN 658 683
CC FT VARIANT 124 124 T -> A (IN STRAIN R6).
CC FT VARIANT 386 388 V -> I (IN STRAIN 63915).
CC FT VARIANT 388 388 D -> E (IN STRAIN R6).
CC FT VARIANT 397 397 E -> K (IN STRAIN 63915).
CC FT VARIANT 523 523 M -> I (IN STRAIN 63915).
CC FT VARIANT 533 533 D -> E (IN STRAIN 45607).
CC FT VARIANT 540 540 T -> S (IN STRAINS 63915 AND R6).
CC FT VARIANT 657 657 N -> S (IN STRAINS 45607 AND R6).
CC SEQUENCE 719 AA; 79758 MW; 5BD397E83B4B3AA6 CRC64;

Query Match 100.0%; Score 666; DB 1; Length 719;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNOLIADLGSRRVNAQANDIPTDLVKAIVSIEDHREFDHGIDITRILGAFLRN 60
 DB 54 KIYDNKNOLIADLGSRRVNAQANDIPTDLVKAIVSIEDHREFDHGIDITRILGAFLRN 113
 QY 61 LOSNSILOGSSTLTQOLIKLTYFSTSDOTISKRAQAMLAIOLEKATIKOELLTYINK 120
 DB 114 LOSNSILOGSSTLTQOLIKLTYFSTSDOTISKRAQAMLAIOLEKATIKOELLTYINK 173
 QY 121 VYMSNNGYVNAQVNAQYVCKDNLNLSLPOLALLAGMPOANOVDPYSHPEAADRRNLV 180
 DB 174 VYMSNNGYVNAQVNAQYVCKDNLNLSLPOLALLAGMPOANOVDPYSHPEAADRRNLV 233
 QY 181 SEKNNOGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEVINOYEETGYNLLT 240
 DB 234 SEKNNOGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEVINOYEETGYNLLT 293
 QY 241 TGMDEVYTNDOEAKHLMIDYNTDEVYANPDELQVASTIVDSNCKVIAQLGARHQQSN 300
 DB 294 TGMDEVYTNDOEAKHLMIDYNTDEVYANPDELQVASTIVDSNCKVIAQLGARHQQSN 353
 QY 301 VSRGIVQAVETNRDWSGTMKPTIDVAPALEYGYDSTATIVHDEPNYNGTNPVYNNMR 360
 DB 354 VSRGIVQAVETNRDWSGTMKPTIDVAPALEYGYDSTATIVHDEPNYNGTNPVYNNMR 413
 QY 361 GFYGNITLQYALQOSRNVAVETLNKVGILNRAKTFPLNGLIDIPSIHYSNAISSNTTESD 420
 DB 414 GFYGNITLQYALQOSRNVAVETLNKVGILNRAKTFPLNGLIDIPSIHYSNAISSNTTESD 473
 QY 421 KKYGASSEKMAAAYAPANGGYTKKPMYTHKVFSDGSEKESNVGTGRAMKETTAYMTD 480
 DB 474 KKYGASSEKMAAAYAPANGGYTKKPMYTHKVFSDGSEKESNVGTGRAMKETTAYMTD 533
 QY 481 MKKTVLYTGGRAYLAMLPOAGKTGTSNTDEIENHITKTSOFVAPDELFACTRYKYSM 540
 DB 534 MKKTVLYTGGRAYLAMLPOAGKTGTSNTDEIENHITKTSOFVAPDELFACTRYKYSM 593
 QY 541 AWMTGYSNRLTPLVNGLTVAAYKVRSMYTLSEGSNPEDMNPEGLYRNGEYVFRNGAR 600
 DB 594 AWMTGYSNRLTPLVNGLTVAAYKVRSMYTLSEGSNPEDMNPEGLYRNGEYVFRNGAR 653
 QY 601 STWNSAPADQPPSTESSSSSSSSSTQSSSTPTSTNNSTTTNNNTQOOSNTTPOOONON 660
 DB 654 STWNSAPADQPPSTESSSSSSSSSTQSSSTPTSTNNSTTTNNNTQOOSNTTPOOONON 713
 QY 661 POPAP 666.
 DB 714 POPAP 719

RESULT 2
 PBPB_STROR STANDARD; PRT: 637 AA.

AC 000573;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Penicillin-binding protein 1A (PBP-1A) (Fragment).
 GN PONA.
 OS Streptococcus oralis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=1303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92325042; PubMed=1624444;
 RA Martin C., Briese T., Hakenbeck R.;
 RT "Nucleotide sequences of genes encoding penicillin-binding proteins
 from Streptococcus pneumoniae and Streptococcus oralis with high
 homology to Escherichia coli penicillin-binding proteins 1a and 1b.";
 RT J. Bacteriol. 174:4517-4523(1992).
 CC -I- FUNCTION: CELL WALL FORMATION.

CC -I- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC
 DR EMBL: M90528; AAA26958.1; -
 DR PIR: B42893; B42893.
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001695; Transglycosyl; 1.
 KW Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
 FT ACT_SITE 371 371
 FT NON_TER 637 637
 FT SEQUENCE 637 AA; 70891 MW; A6D198BCEA603A63 CRC64;
 ACYLATED BY PENICILLIN (BY SIMILARITY).
 SQ

Query Match 8.0%; Score 53; DB 1; Length 637;
 Best Local Similarity 100.0%; Pred. No. 2e-45;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 GKVIAQLGARHQQSNVSEFGINQAVETNRDWSGTMKPTIDVAPALEYGYDSTA 338
 DB 340 GKVIAQLGARHQQSNVSEFGINQAVETNRDWSGTMKPTIDVAPALEYGYDSTA 392

RESULT 3
 PBPB_BUCAI STANDARD; PRT: 760 AA.

ID PBPB_BUCAI
 AC P57296;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Penicillin-binding protein 1B (PBP-1b) (PBP1b) (Murein polymerase)
 DE [Includes: Penicillin-insensitive transglycosylase (EC 2.4.2.-)
 DE (Peptidoglycan Tgase); Penicillin-sensitive transpeptidase
 DE (EC 3.4.-.-) (DD-transpeptidase)].
 GN MRCP OR BU200.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 OS symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TORYO 1998;
 RX MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 RT Buchnera sp. APS.";
 RT Nature 407:81-86(2000).
 RL Nature 407:81-86(2000).
 CC -I- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
 CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
 CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
 CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
 CC TRANSPEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
 CC SUBUNITS) (BY SIMILARITY).
 CC -I- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
 CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC TRANSGLYCOSYLASE FAMILY.
 CC -I- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
 CC TRANSPEPTIDASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AP001118; BAB12917.1;
DR InterPro: IPR001264; Transglycosyl.
DR Pfam: PF00912; Transglycosyl; 1.
DR ProDom: PD001895; Transglycosyl; 1.
KW Peptidoglycan synthetase; Cell wall; Transferase; Glycosyltransferase;
KW Hydrolyase; Multifunctional enzyme; Transmembrane; Signal-anchor;
KW Antibiotic resistance; Complete proteome.
FT DOMAIN 1
FT TRANSMEM 9 29
FT (CYTOPLASMIC (POTENTIAL)).
FT SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT DOMAIN 136 760
FT 392 308
FT 451 684
FT ACT_SITE 451 451
FT ACYLATED BY PENICILLIN (BY SIMILARITY).
SQ SEQUENCE 760 AA; 88002 MW; 00B2C5B51F6947D3 CRC64;

Query Match 1.8%; Score 12; DB 1; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.00098;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 OGSTLTQOLIK 78
DB 204 OGSTLTQOLIK 215
|||||

RESULT 4
MPA_NEUCR STANDARD; PRT: 577 AA.
ID MPA_NEUCR
AC P23955;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitochondrial processing peptidase alpha subunit, mitochondrial
DE precursor (EC 3.4.24.64) (Alpha-MPP).
OS MPP.
GN Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_Taxid=5141;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE OF N-TERMINUS.
RX MEHLIN-90277682; PubMed=2141023;
RA Schneider H., Aretz M., Wachter E., Neupert W.;
RT "Matrix processing peptidase of mitochondria. Structure-function
RT relationships";
RL J. Biol. Chem. 265:9881-9887(1990).
CC -1- FUNCTION: THE MITOCHONDRIAL PROCESSING PROTEASE CLEAVES
CC PRESEQUENCES FROM MITOCHONDRIAL PROTEIN PRECURSORS.
CC -1- CATALYTIC ACTIVITY: Release of N-terminal transit peptides from
CC precursor proteins imported into the mitochondrion, typically with
CC Arg in position P2.
CC -1- SUBUNIT: HETERODIMER OF ALPHA AND BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- DOMAIN: APPEARS TO CONTAIN TWO DOMAINS OF APPROXIMATELY EQUAL SIZE
CC WHICH ARE SEPARATED BY A LOOP-LIKE SEQUENCE.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE
CC INSULINASE FAMILY. DOES NOT SEEM TO HAVE A PROTEASE ACTIVITY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: J05484; AAA33597.1; ALT_SEQ.

DR PIR: A36442; A36442.
DR HSSP: P31800; 1BE3.
DR InterPro: IPR001431; Peptidase_M16.
DR Pfam: PF00675; Peptidase_M16; 1.
DR PROSITE: PS00143; INSULINASE; 1.
KW Hydrolyase; Metalloprotease; Mitochondrion; Transit peptide.
FT TRANSIT 1
FT TRANSIT 35
FT CHAIN 36 577
FT MITOCHONDRIAL
FT ALPHA SUBUNIT. PROCESSING PEPTIDASE
SQ SEQUENCE 577 AA; 63027 MW; 44098F146BBDD1 CRC64;

Query Match 1.5%; Score 10; DB 1; Length 577;
Best Local Similarity 100.0%; Pred. No. 0.081;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 615 ESSSSSSSDS 624
DB 286 ESSSSSSSDS 295
|||||

RESULT 5
PBPB_VIBCH STANDARD; PRT: 777 AA.
ID PBPB_VIBCH
AC Q9KUC0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Penicillin-binding protein 1b (PBP-1b) (PBP1b) (Murein polymerase)
DE (includes: Penicillin-binding protein 1b (PBP-1b) (PBP1b) (Murein polymerase)
DE (Peptidoglycan Gase); Penicillin-sensitive transpeptidase
DE (EC 3.4.2.2) (DD-transpeptidase)).
GN MRCB OR P0NB OR VC0602.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_Taxid=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEHLIN-20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
CC -1- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
CC TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
CC SUBUNTS) (BY SIMILARITY).
CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By
CC similarity).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC TRANSGLYCOSYLASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC TRANSEPTIDASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:46:25 ; Search time 34.32 Seconds
(Without alignments)
3357.069 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 666
1 KIYDNKNQIADLGSERFVN.....TQSNTPDQONQNPQAP 666

Sequence: OLIGO
Gapop 60.0 , Gapext 60.0

Scoring table: Searched: 562222 seqs, 172994929 residues

Word size : 10

Total number of hits satisfying chosen parameters: 81

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP invertebrate: *
6: SP mammal: *
7: SP mhc: *
8: SP organelle: *
9: SP phage: *
10: SP plant: *
11: SP rodent: *
12: SP virus: *
13: SP vertebrate: *
14: SP unclassified: *
15: SP rivirus: *
16: SP bacteriaph: *
17: SP archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	666	100.0	719	2	Q9REU0 streptococc
2	479	71.9	719	2	Q9REU9 streptococc
3	351	52.7	608	2	Q549S1 streptococc
4	351	52.7	719	2	Q9REU3 streptococc
5	274	41.1	398	2	Q9RTM6 streptococc
6	262	39.3	719	2	Q57114 streptococc
7	223	33.5	310	2	Q52739 streptococc
8	223	33.5	310	2	Q52741 streptococc
9	223	33.5	310	2	Q52742 streptococc
10	223	33.5	310	2	Q52743 streptococc
11	207	29.6	718	2	Q9REU7 streptococc
12	192	28.8	398	2	Q87102 streptococc
13	192	27.2	719	2	Q9WU11 streptococc
14	181	27.2	719	2	Q54946 streptococc
15	181	27.2	719	2	Q9WUW0 streptococc
16	181	27.2	719	2	Q9WUW0 streptococc

17	181	27.2	719	2	Q9RET6 streptococc
18	181	27.2	719	2	Q9RET4 streptococc
19	181	27.2	719	2	Q54947 streptococc
20	181	27.2	719	2	Q54948 streptococc
21	181	27.2	719	2	Q54949 streptococc
22	163	24.5	163	2	Q9ETQ2 streptococc
23	142	20.9	163	2	Q9RET8 streptococc
24	139	20.9	163	2	Q9EM43 streptococc
25	134	20.1	310	2	Q52738 streptococc
26	120	18.0	719	2	Q9RET5 streptococc
27	105	15.8	398	2	Q87103 streptococc
28	105	15.8	398	2	Q9RTM7 streptococc
29	105	15.8	398	2	Q87105 streptococc
30	98	14.7	189	2	Q54918 streptococc
31	95	14.3	310	2	Q52737 streptococc
32	85	12.8	398	2	Q9RTM5 streptococc
33	80	12.0	398	2	Q87108 streptococc
34	79	11.9	719	2	Q54950 streptococc
35	76	11.4	398	2	Q87107 streptococc
36	60	9.0	163	2	Q9EM42 streptococc
37	60	9.0	398	2	Q9RTM8 streptococc
38	54	8.1	727	2	Q9ETC4 streptococc
39	53	8.0	398	2	Q87104 streptococc
40	50	7.5	163	2	Q9EM44 streptococc
41	41	6.2	163	2	Q9EM41 streptococc
42	41	6.2	310	2	Q93N72 streptococc
43	38	5.7	527	2	Q70037 streptococc
44	32	4.8	728	2	Q9ETC7 streptococc
45	30	4.5	163	2	Q9EM45 streptococc
46	27	4.1	400	2	Q87106 streptococc
47	23	3.5	721	16	Q99YL1 streptococc
48	22	3.3	163	2	Q9EM40 streptococc
49	20	3.0	139	2	Q52744 streptococc
50	20	3.0	310	2	Q52744 streptococc
51	18	2.7	793	2	Q9EMX9 streptococc
52	14	2.1	664	16	Q9C123 streptococc
53	14	2.1	778	2	Q9EXN1 streptococc
54	11	1.7	196	2	Q66087 streptococc
55	11	1.7	533	2	Q70040 streptococc
56	11	1.7	801	16	Q9C1H4 streptococc
57	11	1.7	821	16	Q97NE4 streptococc
58	11	1.7	821	16	Q97NE4 streptococc
59	11	1.7	966	16	Q9K7X8 streptococc
60	10	1.5	118	5	Q9U1K5 streptococc
61	10	1.5	119	5	Q9V401 streptococc
62	10	1.5	161	4	Q9NW20 streptococc
63	10	1.5	161	4	Q9BVL1 streptococc
64	10	1.5	161	4	Q9NY07 streptococc
65	10	1.5	259	4	Q9H1F6 streptococc
66	10	1.5	261	4	Q9N231 streptococc
67	10	1.5	294	5	Q01796 streptococc
68	10	1.5	504	5	Q02254 streptococc
69	10	1.5	516	2	Q9RQT2 streptococc
70	10	1.5	580	2	Q9K0K9 streptococc
71	10	1.5	584	2	Q9K1J2 streptococc
72	10	1.5	693	2	Q9K1J2 streptococc
73	10	1.5	693	16	Q97N74 streptococc
74	10	1.5	693	2	Q9K1J6 streptococc
75	10	1.5	696	2	Q9K1J2 streptococc
76	10	1.5	714	10	Q9FJ70 streptococc
77	10	1.5	744	16	Q92R95 streptococc
78	10	1.5	755	16	Q98GP0 streptococc
79	10	1.5	767	16	Q92T08 streptococc
80	10	1.5	774	16	Q9X6W0 streptococc
81	10	1.5	775	16	Q9CNH5 streptococc

ALIGNMENTS

RESULT 1
Q9REU0
PRELIMINARY; PRT; 719 AA.

AC Q9REU0;
 DT 01-MAY-2000 (TREMBLrel..13, Created)
 DT 01-MAY-2000 (TREMBLrel..13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel..19, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-SP 1261;
 RX MEDLINE-21432820; PubMed=11549185;
 RA Ferroni A., Berche P.;
 RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
 penicillin-resistant clinical isolates of Streptococcus pneumoniae
 serotype 23F from the nasopharyngeal flora of children."
 RL J. Med. Microbiol. 50:828-832(2001).
 DR EMBL; AF210745; AAF17255.1; -
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79758 MW; 5BD397E83B4B3A6 CRC64;

Query Match 100.0%; Score 666; DB 2; Length 719;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLN 60
 DB 54 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLN 113
 QY 61 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAQEAAMLAIQLEKATKOEILTYINK 120
 DB 114 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAQEAAMLAIQLEKATKOEILTYINK 173
 QY 121 VYMSNGNYGMOTRAQNYKYKDLNNLSLPOLALLAGMPOAPNOYDPSHPEAADRRNLV 180
 DB 174 VYMSNGNYGMOTRAQNYKYKDLNNLSLPOLALLAGMPOAPNOYDPSHPEAADRRNLV 233
 QY 181 SEMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNTLKEYINQVEEETGYNLT 240
 DB 234 SEMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNTLKEYINQVEEETGYNLT 293
 QY 241 TGMADVTTNVDOEAKHLMIDYNTDEVYAVPDELOVASTIVDVSNKVIAQLGARHOSSN 300
 DB 294 TGMADVTTNVDOEAKHLMIDYNTDEVYAVPDELOVASTIVDVSNKVIAQLGARHOSSN 353
 QY 301 VSEFGINQAVETNRDWSSTKPTITDPALEYGVDSTATVHDEPNYNGTNPVYNMNR 360
 DB 354 VSEFGINQAVETNRDWSSTKPTITDPALEYGVDSTATVHDEPNYNGTNPVYNMNR 413
 QY 361 GFEGNTTLOALQOQRNVAVETLNKVGINRAKTFNLGIGIDYPSIHYSNALSSNTTESD 420
 DB 414 GFEGNTTLOALQOQRNVAVETLNKVGINRAKTFNLGIGIDYPSIHYSNALSSNTTESD 473
 QY 421 KRYGASSEKMAAAYAAFANGGTYKPMYIHKVYVSDGSEKESNVGTRAMKETTAAYMMT 480
 DB 474 KRYGASSEKMAAAYAAFANGGTYKPMYIHKVYVSDGSEKESNVGTRAMKETTAAYMMT 533
 QY 481 MKKTYVITGTGRNAYLAWLPOAGKTGTSNYTDDEITENHKTQSFAVPDELFAGYTKRYKM 540
 DB 534 MKKTYVITGTGRNAYLAWLPOAGKTGTSNYTDDEITENHKTQSFAVPDELFAGYTKRYKM 593
 QY 541 AYWGTGSNLTPLVGNGLVAAKVRSMATYISEGSPEDDMNIPELLYRNGEVEFNKGAR 600
 DB 594 AYWGTGSNLTPLVGNGLVAAKVRSMATYISEGSPEDDMNIPELLYRNGEVEFNKGAR 653
 QY 601 STWNSPAPQPPSTESSSSSDSSSTPSQSSSTPTSTNNSTTTNPNNNTTQOOSNTTPOOONON 660

DB 654 STWNSPAPQPPSTESSSSSDSSSTPSQSSSTPTSTNNSTTTNPNNNTTQOOSNTTPOOONON 713
 QY 661 POPAP 666
 DB 714 POPAP 719

RESULT 2

ID Q9RET9 PRELIMINARY; PRT; 719 AA.

AC Q9RET9;
 DT 01-MAY-2000 (TREMBLrel..13, Created)
 DT 01-MAY-2000 (TREMBLrel..13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel..19, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-SP 1513;
 RX MEDLINE-21432820; PubMed=11549185;
 RA Ferroni A., Berche P.;
 RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
 penicillin-resistant clinical isolates of Streptococcus pneumoniae
 serotype 23F from the nasopharyngeal flora of children."
 RL J. Med. Microbiol. 50:828-832(2001).
 DR EMBL; AF210746; AAF17256.1; -
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79830 MW; 5F6776B8DEAE3840 CRC64;

Query Match 71.9%; Score 479; DB 2; Length 719;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLN 60
 DB 54 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLN 113
 QY 61 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAQEAAMLAIQLEKATKOEILTYINK 120
 DB 114 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAQEAAMLAIQLEKATKOEILTYINK 173
 QY 121 VYMSNGNYGMOTRAQNYKYKDLNNLSLPOLALLAGMPOAPNOYDPSHPEAADRRNLV 180
 DB 174 VYMSNGNYGMOTRAQNYKYKDLNNLSLPOLALLAGMPOAPNOYDPSHPEAADRRNLV 233
 QY 181 SEMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNTLKEYINQVEEETGYNLT 240
 DB 234 SEMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNTLKEYINQVEEETGYNLT 293
 QY 241 TGMADVTTNVDOEAKHLMIDYNTDEVYAVPDELOVASTIVDVSNKVIAQLGARHOSSN 300
 DB 294 TGMADVTTNVDOEAKHLMIDYNTDEVYAVPDELOVASTIVDVSNKVIAQLGARHOSSN 353
 QY 301 VSEFGINQAVETNRDWSSTKPTITDPALEYGVDSTATVHDEPNYNGTNPVYNMNR 360
 DB 354 VSEFGINQAVETNRDWSSTKPTITDPALEYGVDSTATVHDEPNYNGTNPVYNMNR 413
 QY 361 GFEGNTTLOALQOQRNVAVETLNKVGINRAKTFNLGIGIDYPSIHYSNALSSNTTESD 420
 DB 414 GFEGNTTLOALQOQRNVAVETLNKVGINRAKTFNLGIGIDYPSIHYSNALSSNTTESD 473
 QY 421 KRYGASSEKMAAAYAAFANGGTYKPMYIHKVYVSDGSEKESNVGTRAMKETTAAYMMT 479

Db 474 KKGASSEKMAAAYAFANGGTYKKPMYHKVPSDGESEKFSNVGTRAMKETTAYMMT 532

RESULT 3
ID 054951 PRELIMINARY; PRT: 608 AA.
AC 054951;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN Streptococcus pneumoniae.
OS Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8250;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-resistant Streptococcus pneumoniae isolated in
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain.";
RL EMBO J. 11:3831-3836(1992).
DR EMBL; X67871; CAA48071.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam; PF00912; Transglycosyl; 1.
DR Pfam; PF00905; Transpeptidase; 1.
DR ProDom; PD001895; Transglycosyl; 1.
FT NON_TER 608
SQ SEQUENCE 608 AA; 67918 MW; 68F5C09E2AF9114 CRC64;

Query Match 52.7%; Score 351; DB 2; Length 608;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNLIDLGSERRVNAQANDIPDYLKAIYSIEDHREFDRGIDTIRILGAFIRN 60
DB 54 KIYDNKNLIDLGSERRVNAQANDIPDYLKAIYSIEDHREFDRGIDTIRILGAFIRN 113
QY 61 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSIRKQAEAMLAIQLEOKATKQEIILTYINK 120
DB 114 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSIRKQAEAMLAIQLEOKATKQEIILTYINK 173
QY 121 YVMSNGNTGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPPEAQRRLNVL 180
DB 174 YVMSNGNTGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPPEAQRRLNVL 233
QY 181 SEMKNQGYISAEQYKAVNPTDGLQSLKSASNPAYMDNVLKENVINOVEERTGYNLLT 240
DB 234 SEMKNQGYISAEQYKAVNPTDGLQSLKSASNPAYMDNVLKENVINOVEERTGYNLLT 293
QY 241 TGMVYTNVDOEAQKHLMDIYNTDEVYVAPDDELQVASTIYDVNSGKYIAQLGARHSSN 300
DB 294 TGMVYTNVDOEAQKHLMDIYNTDEVYVAPDDELQVASTIYDVNSGKYIAQLGARHSSN 353
QY 301 VSFGINQAVETNRDWSGTMKPTTDYAPALEGVYDSTATIYHDEPYNPGT 351
DB 354 VSFGINQAVETNRDWSGTMKPTTDYAPALEGVYDSTATIYHDEPYNPGT 404

RESULT 4
Q9RET3 PRELIMINARY; PRT: 719 AA.
ID Q9RET3;
AC Q9RET3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN BPPIA.

OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BM 4200;
RX MEDLINE=21432820; PubMed=11549185;
RA Ferrioli A., Berche P.;
RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
RT penicillin-resistant clinical isolates of Streptococcus pneumoniae
RT serotype 23F from the nasopharyngeal flora of children.";
RL J. Med. Microbiol. 50:828-832(2001).
DR EMBL; AF210752; AAF17262.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam; PF00912; Transglycosyl; 1.
DR Pfam; PF00905; Transpeptidase; 1.
DR ProDom; PD001895; Transglycosyl; 1.
SQ SEQUENCE 719 AA; 79781 MW; DB5993F6BD47F72D CRC64;

Query Match 52.7%; Score 351; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNLIDLGSERRVNAQANDIPDYLKAIYSIEDHREFDRGIDTIRILGAFIRN 60
DB 54 KIYDNKNLIDLGSERRVNAQANDIPDYLKAIYSIEDHREFDRGIDTIRILGAFIRN 113
QY 61 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSIRKQAEAMLAIQLEOKATKQEIILTYINK 120
DB 114 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSIRKQAEAMLAIQLEOKATKQEIILTYINK 173
QY 121 YVMSNGNTGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPPEAQRRLNVL 180
DB 174 YVMSNGNTGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPPEAQRRLNVL 233
QY 181 SEMKNQGYISAEQYKAVNPTDGLQSLKSASNPAYMDNVLKENVINOVEERTGYNLLT 240
DB 234 SEMKNQGYISAEQYKAVNPTDGLQSLKSASNPAYMDNVLKENVINOVEERTGYNLLT 293
QY 241 TGMVYTNVDOEAQKHLMDIYNTDEVYVAPDDELQVASTIYDVNSGKYIAQLGARHSSN 300
DB 294 TGMVYTNVDOEAQKHLMDIYNTDEVYVAPDDELQVASTIYDVNSGKYIAQLGARHSSN 353
QY 301 VSFGINQAVETNRDWSGTMKPTTDYAPALEGVYDSTATIYHDEPYNPGT 351
DB 354 VSFGINQAVETNRDWSGTMKPTTDYAPALEGVYDSTATIYHDEPYNPGT 404

RESULT 5
Q9R7M6 PRELIMINARY; PRT: 398 AA.
ID Q9R7M6;
AC Q9R7M6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN BPPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=10/219;
RX MEDLINE=98409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a Thr-371 substitution in a conserved amino acid motif
RT of penicillin-binding protein 1A with penicillin resistance of
RT Streptococcus pneumoniae.";
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).

DR EMBL: AB006874; BAA32070.1; -
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00905; Transpeptidase; 1.
 FT NON_TER 1
 FT NON_TER 398
 SQ SEQUENCE 398 AA; 43613 MW; D2C0E784E9B3451 CRC64;

Query Match 41.1%; Score 274; DB 2; Length 398;
 Best Local Similarity 99.7%; Pred. No. 3.9e-270;
 Matches 374; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 267 VAYPDELQAVASTIVDVNSGKVTIAOLGARHOSNSVSGINQAVETNRDMSGTMRKPTTDTYA 326
 DB 1 VAYPDELQAVASTIVDVNSGKVTIAOLGARHOSNSVSGINQAVETNRDMSGTMRKPTTDTYA 60
 QY 327 PALEGVYSTATIVHDEPVNPGTNPVYNDMGRGFGNITTOYALQOQRNPAVETLNK 386
 DB 61 PALEGVYSTATIVHDEPVNPGTNPVYNDMGRGFGNITTOYALQOQRNPAVETLNK 120
 QY 387 VGLNRAKTEFLNGIDYPSIHYSNATSSNTTESDKKYGASSEKMAAAYAFANGCTYKRP 446
 DB 121 VGLNRAKTEFLNGIDYPSIHYSNATSSNTTESDKKYGASSEKMAAAYAFANGCTYKRP 180
 QY 447 MTIHKVYFSDGSEKERSNVGTGRAMKETAYMMTDMKTYLVTGTGNAYLAMPQAGKTG 506
 DB 181 MTIHKVYFSDGSEKERSNVGTGRAMKETAYMMTDMKTYLVTGTGNAYLAMPQAGKTG 240
 QY 507 TSNYTDDEETENHIKTQFOFAPDELFGYTRKYSMAVWTGYSNLTPLVGNGLVAAKYR 566
 DB 241 TSNYTDDEETENHIKTQFOFAPDELFGYTRKYSMAVWTGYSNLTPLVGNGLVAAKYR 300
 QY 567 SMWYTLSEGSNPEDMNIPEGLYRNGEFVRKNGARSTWNSPAPQPPSTSSSSSDSSTS 626
 DB 301 SMWYTLSEGSNPEDMNIPEGLYRNGEFVRKNGARSTWNSPAPQPPSTSSSSSDSSTS 360
 QY 627 QSSSTPTSTNNSTTT 641
 DB 361 QSSSTPTSTNNSTTT 375

RESULT 6
 ID 057114 PRELIMINARY; PRT; 719 AA.
 AC 057114;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN 111
 RP SEQUENCE FROM N.A.
 RC STRAIN=C5111 (CEPHALOSPORIN-RESISTANT CLINICAL ISOLATE);
 RX MEDLINE=96012191; PubMed=7574521;
 RA Coffey T.J., Daniels M., McDougal L.K., Dowson C.G., Tenover F.C., Spratt B.G.;
 RA "Genetic analysis of clinical isolates of Streptococcus pneumoniae with high-level resistance to expanded-spectrum cephalosporins";
 RL Antimicrob. Agents Chemother. 39:1306-1313(1995).
 DR EMBL: 249095; CAA88918.1; -
 DR EMBL: 249094; CAA88917.1; -
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79712 MW; 5396607C0E67D06E CRC64;

Query Match 39.3%; Score 262; DB 2; Length 719;
 Best Local Similarity 100.0%; Pred. No. 1.1e-257;
 Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KITDKNKQLIADIGSEFRVNAQANDIPTDLVKAVISIEHFRFDHNGIDTIRILGAFNR 60
 DB 54 KITDKNKQLIADIGSEFRVNAQANDIPTDLVKAVISIEHFRFDHNGIDTIRILGAFNR 113
 QY 61 LOSNSLOGSTTLQOILIKLYFTSTSDOTISRAQOEWALAIQEQATKQOILTYINK 120
 DB 114 LOSNSLOGSTTLQOILIKLYFTSTSDOTISRAQOEWALAIQEQATKQOILTYINK 173
 QY 121 VYMSNGYGMQTAQNVYCKDLNNLSLPLLALLGAPQAPNOVDYPSHPAADRNLV 180
 DB 174 VYMSNGYGMQTAQNVYCKDLNNLSLPLLALLGAPQAPNOVDYPSHPAADRNLV 233
 QY 181 SEKNNGYISAEQYERAVNPTIDGQSLSASNTAYMDNYLKEYINOVEETGYNLT 240
 DB 234 SEKNNGYISAEQYERAVNPTIDGQSLSASNTAYMDNYLKEYINOVEETGYNLT 293
 QY 241 TGMVYTNVDOEAQKHLMDIYN 262
 DB 294 TGMVYTNVDOEAQKHLMDIYN 315

RESULT 7
 ID 052739 PRELIMINARY; PRT; 310 AA.
 AC 052739;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
 GN PONA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN 111
 RP SEQUENCE FROM N.A.
 RC STRAIN=64429;
 RX MEDLINE=98287565; PubMed=9624469;
 RA Smith A.M., Klugman K.P.;
 RA "Alterations in PBPIA essential for high-level penicillin resistance in Streptococcus pneumoniae";
 RL Antimicrob. Agents Chemother. 42:1329-1333(1998).
 DR EMBL: AF046233; AAC24698.1; -
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00905; Transpeptidase; 1.
 FT NON_TER 1
 FT NON_TER 310
 SQ SEQUENCE 310 AA; 34445 MW; 31C692D48789576D CRC64;

Query Match 33.5%; Score 223; DB 2; Length 310;
 Best Local Similarity 100.0%; Pred. No. 2.7e-218;
 Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 LMDIYITDEYVAYPDELQAVASTIVDVNSGKVTIAOLGARHOSNSVSGINQAVETNRDMSG 316
 DB 1 LMDIYITDEYVAYPDELQAVASTIVDVNSGKVTIAOLGARHOSNSVSGINQAVETNRDMSG 60
 QY 317 STMKPTTDAAPALEGVYSTATIVHDEPVNPGTNPVYNDMGRGFGNITTOYALQOQR 376
 DB 61 STMKPTTDAAPALEGVYSTATIVHDEPVNPGTNPVYNDMGRGFGNITTOYALQOQR 120
 QY 377 NVPAYETLNKVGINRAKTEFLNGIDYPSIHYSNATSSNTTESDKKYGASSEKMAAAYAA 436
 DB 121 NVPAYETLNKVGINRAKTEFLNGIDYPSIHYSNATSSNTTESDKKYGASSEKMAAAYAA 180
 QY 437 FANGCTYKRPMTIHKVYFSDGSEKERSNVGTGRAMKETAYMMT 479
 DB 181 FANGCTYKRPMTIHKVYFSDGSEKERSNVGTGRAMKETAYMMT 223

RESULT	8			
052741				
ID	052741	PRELIMINARY;	PFT;	310 AA.
AC	052741;			
DT	01-JUN-1998	(TREMBREL. 06, Created)		
DT	01-JUN-1998	(TREMBREL. 06, Last sequence update)		
DT	01-DEC-2001	(TREMBREL. 19, Last annotation update)		
DE	PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).			
GN	PONA.			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
CC	Streptococcus.			
OX	NCBL_Taxid=1313;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=7851;			
RX	MEDLINE=98287565; PubMed=9624469;			
RA	Smith A.M., Klugman K.P.:			
RT	"Alterations in PBP 1A essential for high-level penicillin resistance			
RT	in Streptococcus pneumoniae."			
RL	Antimicrob. Agents Chemother. 42:1329-1333(1998).			
DR	EMBL: AF046235; AAC24700.1; -			
DR	InterPro: IPR001460; Transceptidse.			
DR	Pfam: PF00905; Transpeptidase; 1.			
FT	NON_TER	1		
FT	NON_TER	310		
SEQ	SEQUENCE	310 AA; 34433 MW; 6E2BDEFA50AB28CF3 CRC64;		

Query Match	33.5%	Score 223	DB 2:	Length 310
Best Local Similarity	100.0%	Pred. NO.	2.7e-218	
Matches 223	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy	257	LMIYNTDEVAAYPDDLEJOVASTIVDSNCKVLIAOLGARHSSNSNGICINOAETNRDMG	316
Dd	1	LMIDYNDEVAIYPDDELJOVASTIVDSNCKVLIAQLGARHSSSVSGINOAAYETNRDMG	60
Qy	317	STMRPTDYAPALEGYGYDSTATIVHDPEPNYPGTNPVVNMGRGEYGNITLLOYALQOSR	376
Dd	61	STMRPTIDYAPALEGYGYDSTATIVHDPEPNYPGTNPVVNMGRGYGNITLLOYALQOSR	120
Qy	377	NVPAVETLNKVGELNRAKFTFLNGIGIDPSIHYSNAISSNTTESDKKYAGASSEKMAAAYA	436
Dd	121	NVPAVETLNKVGELNRAKFTFLNGIGIDPSIHYSNAISSNTTESDKKYAGASSEKMAAAYA	180
Qy	437	FANGCTYYKKPYTHKVYESDGSEKEFSNVGTRAMKETTAIYMT	479
Dd	181	FANGCTYYKKPYTHKVYESDGSEKEFSNVGTRAMKETTAIYMT	223

RESULT	9			
052742				
ID	052742	PRELIMINARY;	PRT;	310 AA.
AC	052742;			
DT	01-JUN-1998	(TREMBLrel. 06, Created)		
DT	01-JUN-1998	(TREMBLrel. 06, last sequence update)		
DT	01-DEC-2001	(TREMBLrel. 19, last annotation update)		
DE	PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).			
GN	PONA.			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1313;			
RN	{1}			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NG4;			
RX	MEDLINE=98287565; Pubmed=9624469;			
RA	Smith A.M., Klugman K.P.:			
RT	"Alterations in PBP 1A essential for high-level penicillin resistance			
RL	in Streptococcus pneumoniae.";			
TL	Antimicrob. Agents Chemother. 42:1329-1333(1998).			

DR	EMBL: AF046236; AAC24701.1; -;
DR	InterPro: IPR001460; Transseptidase.
DR	Pfam: PF00905; Transseptidase; 1.
FT	NON_TER 1
FT	NON_TER 310 310
SQ	SEQUENCE 310 AA: 34480 MW: 3964987387C5A56D CRC64;

Query Match	33.5%	Score	223	DB	2	Length	310
Best Local Similarity	100.0%	Pred	No.	2.7e-218			
Matches	223	Conservative	0	Mismatches	0	Indels	0
						Gaps	0

Oy	257	LMIDINTBEYVYPPDELOVASTIYDVNSGKYIAOLGAHOSNNVSFGINQAVETNRDMG	316
Db	1	LMIDINTBEYVYPPDELOVASTIYDVNSGKYIAOLGAHOSNNVSFGINQAVETNRDMG	60
Oy	317	STMKEITDYAPALEYGVYDSTATIYHDEPYNPFGTNTPYVNMDDRGYFGNITLOVALOOSR	376
Db	61	STMKEITDYAPALEYGVYDSTATIYHDEPYNPFGTNTPYVNMDDRGYFGNITLOVALOOSR	120
Oy	377	NVPVETLNKVGVLNRAKFTPLNGLGIDYPSIYHSNAISSNTTESDKKYGASSEKMAAAYAA	436
Db	121	NVPVETLNKVGVLNRAKFTPLNGLGIDYPSIYHSNAISSNTTESDKKYGASSEKMAAAYAA	180
Oy	437	FANGGTYYKPMYIHKVVPFSDGSEKEFSNNGTAMKETTAYMMT	479
Db	181	FANGGTYYKPMYIHKVVPFSDGSEKEFSNNGTAMKETTAYMMT	223

RESULT 10

ID	052743	PRELIMINARY;	PRT:	310 AA.
AC	052743;			
DT	01-JUN-1998 (TREMBLrel. 06, Created)			
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).			
GN	PONA.			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Streptococcus.			
OX	NCHI_TaxID=1313;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=I7619;			
RX	MEDLINE=98287565; PubMed=9624469;			
RA	Smith A.M., Klugman K.P.;			
RT	"Alterations in PBP 1A essential for high-level penicillin resistance			
RT	in Streptococcus pneumoniae."			
RL	Antimicrob. Agents Chemother. 42:1329-1333(1998).			
DR	EMBL; AF046237; AAC24702.1; -			
DR	InterPro: IPR001460; Transpeptidase.			
DR	Pfam: PF00905; Transpeptidase; 1.			
FT	NON_TER 1			
FT	NON_TER 310			
SEQUENCE	310 AA: 34461 MW: 6628021607528CF3 CRC64;			

Query Match	33.5%	Score 223	DB 2	Length 310
Best Local Similarity	100.0%	Pred. No. 2	7e-218	
Matches 223; Conservative	0	Mismatches 0	Indels 0	Gaps 0

QY	237	LMDIYNDEIYAAVPDEELOVASTIVDVNSGKIAIOLGARHOSSNVSFCINQVAETNRDMG	316
Db	1	LMDIYNFDEIYAAVPDEELOVASTIVDVNSGKIAIOLGARHOSSNVSFCINQVAETNRDMG	60
QY	317	STMKEITDYAPALEYGVYDSTATIYHDEPYNPGINTPIYMMDRGYFGNITTLQYALOOSR	376
Db	61	STMKEITDYAPALEYGVYDSTATIYHDERYNPCTNTPYMMDRGYFGNITTLQYALOOSR	120
QY	377	NVPAVEITLNTKVGILNNAKTFNLNLGIDIPSIHNSNAISSNTTSSDKKYGASSSEKMAAAVYA	436
Db	121	NVPAVEITLNTKVGILNNAKTFNLNLGIDIPSIHNSNAISSNTTSSDKKYGASSSEKMAAAVYA	180

OY 437 FANGGTYRKPMYIHKVFSDSSEKFSNVGTRAMKETAYMMT 479
 DB 181 FANGGTYRKPMYIHKVFSDSSEKFSNVGTRAMKETAYMMT 223

RESULT 11

052740 PRELIMINARY; PRT; 310 AA.
 ID 052740;
 AC 052740;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 RT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
 GN PONA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56739;
 RX MEDLINE=98287565; PubMed=9624469;
 RA Smith A.M., Klugman K.P.;
 RT "Alterations in Pbp 1A essential for high-level penicillin resistance
 in Streptococcus pneumoniae."
 RL Antimicrob. Agents Chemother. 42:1329-1333(1998).
 DR EMBL: AF046234; AAC24699.1; -
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00905; Transpeptidase; 1.
 FT NON_TER 1 310
 FT NON_TER 1
 SQ SEQUENCE 310 AA; 34418 MW; 396418EF9C15A763 CRC64;

Query Match 31.1%; Score 207; DB 2; Length 310;
 Best Local Similarity 100.0%; Pred. No. 5.4e-202;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 257 LMDIYNTDEYVAPPDELOVASTIVDSNGKYIAQLGARHSSNVSEFGINQAVETNRDMG 316
 DB 1 LMDIYNTDEYVAPPDELOVASTIVDSNGKYIAQLGARHSSNVSEFGINQAVETNRDMG 60
 OY 317 STMKPIITDYAPALEYGVYSTATIVHDEPYNPGTNPVYNDMDRGYFNTLQYALQGR 376
 DB 61 STMKPIITDYAPALEYGVYSTATIVHDEPYNPGTNPVYNDMDRGYFNTLQYALQGR 120
 OY 377 NPPAVETLNKVGINRAKTFPLNGIDYPSIHSNAISSMTTESDCKRYGASSEKMAAAYAA 436
 DB 121 NPPAVETLNKVGINRAKTFPLNGIDYPSIHSNAISSMTTESDCKRYGASSEKMAAAYAA 180
 OY 437 FANGGTYRKPMYIHKVFSDSSEKFS 463
 DB 181 FANGGTYRKPMYIHKVFSDSSEKFS 207

RESULT 12

09RET7 PRELIMINARY; PRT; 718 AA.
 ID 09RET7;
 AC 09RET7;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
 GN PBp1A.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SP 1258;
 RX MEDLINE=21432820; PubMed=11549185;

RA Ferroni A., Berche P.;
 RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
 penicillin-resistant clinical isolates of Streptococcus pneumoniae
 RT serotype 23F from the nasopharyngeal flora of children."
 RL J. Med. Microbiol. 50:828-832(2001).

DR EMBL: AF210748; AAF17258.1; -
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR Prodom: PD001895; Transglycosyl; 1.
 FT NON_TER 718 718
 SQ SEQUENCE 718 AA; 79715 MW; 074216E49616C3E4 CRC64;

Query Match 29.6%; Score 197; DB 2; Length 718;
 Best Local Similarity 100.0%; Pred. No. 1.8e-191;
 Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KIYDNKNQIADISSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLENN 60
 DB 54 KIYDNKNQIADISSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLENN 113
 OY 61 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKRAQEAAMLAIQLEKATKQELITYINK 120
 DB 114 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKRAQEAAMLAIQLEKATKQELITYINK 173
 OY 121 VVMSNGNMGQTAQONYVYGKDLNNLSLPOLALLAGMPQAPNOVDPSHPEDAADRNNLV 180
 DB 174 VVMSNGNMGQTAQONYVYGKDLNNLSLPOLALLAGMPQAPNOVDPSHPEDAADRNNLV 233
 OY 181 SEMKNOGYISAQYERKA 197
 DB 234 SEMKNOGYISAQYERKA 250

RESULT 13
 087102 PRELIMINARY; PRT; 398 AA.
 ID 087102;
 AC 087102;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
 GN PBp1A.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H1/H23;
 RX MEDLINE=98409715; PubMed=9736547;
 RA Asahi Y., Ubukata K.;
 RT "Association of a Thr-371 substitution in a conserved amino acid motif
 of penicillin-binding protein 1A with penicillin resistance of
 Streptococcus pneumoniae."
 RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
 DR EMBL: AB006868; BAA32064.1; -
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00905; Transpeptidase; 1.
 FT NON_TER 398 398
 FT NON_TER 1
 SQ SEQUENCE 398 AA; 43584 MW; 47EA32CA235276AE CRC64;

Query Match 28.8%; Score 192; DB 2; Length 398;
 Best Local Similarity 100.0%; Pred. No. 1.3e-186;
 Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 267 VAPPDELOVASTIVDSNGKYIAQLGARHSSNVSEFGINQAVETNRDMGSTMKPIITDYA 326
 DB 1 VAPPDELOVASTIVDSNGKYIAQLGARHSSNVSEFGINQAVETNRDMGSTMKPIITDYA 60


```

QY 327 PALEGYVDSTATIVHDEPNYPGTNTPVYNMNRGFCGNTLOYALQOSRNPVAVETLTK 386
Db 61 PALEGYVDSTATIVHDEPNYPGTNTPVYNMNRGFCGNTLOYALQOSRNPVAVETLTK 120
QY 387 VGLNRKATFLNGLGIDIPYSIHYSNAISSNTFESDKKYGASSEKMAAAYAAPANGCTYYKP 446
Db 121 VGLNRKATFLNGLGIDIPYSIHYSNAISSNTFESDKKYGASSEKMAAAYAAPANGCTYYKP 180
QY 447 MYHKVVSDDGS 458
Db 181 MYHKVVSDDGS 192

RESULT 14
09WM11 PRELIMINARY: PRT: 719 AA.
AC 09WM11;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN BPPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=URU-E159, M134, URU-E135, AND URU-E157;
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
RT clone of Streptococcus pneumoniae arose by large recombinational
RT replacements of the cpsA-pbpA region.";
RL Microbiology 0:0-0(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=URU-E206;
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
RT clone of Streptococcus pneumoniae arose by large recombinational
RT replacements of the cpsA-pbpA region.";
RL Microbiology 145:0-0(1999).
DR EMBL: AF139890; AAD43073.1; -.
DR EMBL: AF139887; AAD43070.1; -.
DR EMBL: AF139888; AAD43071.1; -.
DR EMBL: AF139889; AAD43072.1; -.
DR EMBL: AF159448; AAD48430.1; -.
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl. 1.
DR Pfam: PF00905; Transpeptidase. 1.
DR ProDom: PD001895; Transglycosyl. 1.
SQ SEQUENCE 719 AA; 79701 MW; 3EDC8A21C9ADAC97 CRC64;

Query Match 27.2%; Score 181; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 3.6e-175;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 IRIIGAFLRNLSNSLOGSSTLTQOLIKLTFYSTSDQTSRKAEAMLAIOLEOKATK 110
Db 104 IRIIGAFLRNLSNSLOGSSTLTQOLIKLTFYSTSDQTSRKAEAMLAIOLEOKATK 163
QY 111 QEILTYINKYMSNGNTAGMOTAQNYGKDLNLSLPQALLAGMPQAPQYDPYSHE 170
Db 164 QEILTYINKYMSNGNTAGMOTAQNYGKDLNLSLPQALLAGMPQAPQYDPYSHE 223
QY 171 AAQDRRLNVISEMKNQGTISAEQYKAVNPITDGLQSLKASNPAYMDNYLKEVINOV 230
Db 224 AAQDRRLNVISEMKNQGTISAEQYKAVNPITDGLQSLKASNPAYMDNYLKEVINOV 283
QY 231 E 231

```

```

Db 284 E 284

RESULT 15
Q54946 PRELIMINARY: PRT: 719 AA.
ID Q54946;
AC Q54946;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA OR BPPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=456;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain.";
RL EMBO J. 11:3831-3836(1992).
RN [2]
RP SEQUENCE OF 310-619 FROM N.A.
RC STRAIN=8303, AND 35193;
RA Smith A.M., Klugman K.P.;
RL submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 320-717 FROM N.A.
RC STRAIN=#17/246, #20/B99, AND #27/SHA3;
RX MEDLINE=96409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a thr-371 substitution in a conserved amino acid motif
RT of penicillin-binding protein 1A with penicillin resistance of
RT Streptococcus pneumoniae.";
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=SP-665;
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
RT clone of Streptococcus pneumoniae arose by large recombinational
RT replacements of the cpsA-pbpA region.";
RL Microbiology 0:0-0(1999).
DR EMBL: X67868; CAA48068.1; -.
DR EMBL: AF046230; AAC24695.1; -.
DR EMBL: AB006876; BAA32072.1; -.
DR EMBL: AF139883; AAD43066.1; -.
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl. 1.
DR Pfam: PF00905; Transpeptidase. 1.
DR ProDom: PD001895; Transglycosyl. 1.
FT NON_TER 719
SQ SEQUENCE 719 AA; 79736 MW; F090F91522C834DC CRC64;

Query Match 27.2%; Score 181; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 3.6e-175;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 IRIIGAFLRNLSNSLOGSSTLTQOLIKLTFYSTSDQTSRKAEAMLAIOLEOKATK 110
Db 104 IRIIGAFLRNLSNSLOGSSTLTQOLIKLTFYSTSDQTSRKAEAMLAIOLEOKATK 163
QY 111 QEILTYINKYMSNGNTAGMOTAQNYGKDLNLSLPQALLAGMPQAPQYDPYSHE 170
Db 164 QEILTYINKYMSNGNTAGMOTAQNYGKDLNLSLPQALLAGMPQAPQYDPYSHE 223

```


QY 171 AAODRRNLVLESEKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNYLKEVINOV 230
 DB 224 AAODRRNLVLESEKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNYLKEVINOV 283
 QY 231 E 231
 DB 284 E 284

RESULT 16

Q9WVW0 PRELIMINARY; PRT: 719 AA.
 AC Q9WVW0; 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBP1A.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PO-342, PO-273, AND PO-341;
 RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
 RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
 RT clone of Streptococcus pneumoniae arose by large recombinational
 RT replacements of the cpsA-pbpA region."
 RL Microbiology 0:0-0(1999).
 DR EMBL; AF139886; AADA3069.1; -;
 DR EMBL; AF139884; AADA3067.1; -;
 DR EMBL; AF139885; AADA3068.1; -;
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79622 MW; 7FBA6A75EA8FF8B3 CRC64;

Query Match 27.2%; Score 181; DB 2; Length 719;
 Best Local Similarity 100.0%; Pred. No. 3.6e-175;
 Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 IRLGAFLENLQSNLSQGSSTLTQOLIKLTFSTSTSDQTSRKAQEWLAIOLEQKATK 110
 DB 104 IRLGAFLENLQSNLSQGSSTLTQOLIKLTFSTSTSDQTSRKAQEWLAIOLEQKATK 163
 QY 111 QELITYYINKVYMSNGNYGMQTAQNYGKDLNNLSLPQALLAGMPQAPNOYDPYSHP 170
 DB 164 QELITYYINKVYMSNGNYGMQTAQNYGKDLNNLSLPQALLAGMPQAPNOYDPYSHP 223
 QY 171 AAODRRNLVLESEKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNYLKEVINOV 230
 DB 224 AAODRRNLVLESEKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNYLKEVINOV 283
 QY 231 E 231
 DB 284 E 284

RESULT 17

Q9RET6 PRELIMINARY; PRT: 719 AA.
 AC Q9RET6; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBP1A.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SP 1053; PubMed=11549185;
 RX MEDLINE=21432820; PubMed=11549185;
 RA Ferroni A., Berche P.;
 RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
 RT penicillin-resistant clinical isolates of Streptococcus pneumoniae
 RT serotype 23F from the nasopharyngeal flora of children."
 RL J. Med. Microbiol. 50:828-832(2001).
 DR EMBL; AF210749; AAF17259.1; -;
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79638 MW; 1DAE3C5937048E16 CRC64;

Query Match 27.2%; Score 181; DB 2; Length 719;
 Best Local Similarity 100.0%; Pred. No. 3.6e-175;
 Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 IRLGAFLENLQSNLSQGSSTLTQOLIKLTFSTSTSDQTSRKAQEWLAIOLEQKATK 110
 DB 104 IRLGAFLENLQSNLSQGSSTLTQOLIKLTFSTSTSDQTSRKAQEWLAIOLEQKATK 163
 QY 111 QELITYYINKVYMSNGNYGMQTAQNYGKDLNNLSLPQALLAGMPQAPNOYDPYSHP 170
 DB 164 QELITYYINKVYMSNGNYGMQTAQNYGKDLNNLSLPQALLAGMPQAPNOYDPYSHP 223
 QY 171 AAODRRNLVLESEKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNYLKEVINOV 230
 DB 224 AAODRRNLVLESEKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNYLKEVINOV 283
 QY 231 E 231
 DB 284 E 284

RESULT 18

Q9RET4 PRELIMINARY; PRT: 719 AA.
 AC Q9RET4; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBP1A.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SP 22861;
 RX MEDLINE=21432820; PubMed=11549185;
 RA Ferroni A., Berche P.;
 RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
 RT penicillin-resistant clinical isolates of Streptococcus pneumoniae
 RT serotype 23F from the nasopharyngeal flora of children."
 RL J. Med. Microbiol. 50:828-832(2001).
 DR EMBL; AF210751; AAF17261.1; -;
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79567 MW; 7B6EDCCBB8BCF286 CRC64;

Query Match 27.2%; Score 181; DB 2; Length 719;

Best Local Similarity 100.0%; Pred. No. 3.6e-175;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 51 IRLGAFRLNLSNLSGGSTLTQOLIKITFTSTSDOTISRKAOEAWLAIOLEOKATK 110
DB 104 IRLGAFRLNLSNLSGGSTLTQOLIKITFTSTSDOTISRKAOEAWLAIOLEOKATK 163
OY 111 QEILFYINKVYMSNGNGMGTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHE 170
DB 164 QEILFYINKVYMSNGNGMGTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHE 223
OY 171 AAODRRNLVLESMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINOV 230
DB 224 AAODRRNLVLESMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINOV 283
OY 231 E 231
DB 284 E 284

RESULT 19
O54947 PRELIMINARY; PRT: 719 AA.
AC O54947;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DF 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=681;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain."
RL EMBL J. 11:3831-3836(1992).
DR EMBL; X67867; CAA48067.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
FT NON_TER 719
SQ SEQUENCE 719 AA; 79710 MW; D90EB42D5380D9E1 CRC64;
```

Query Match 27.2%; Score 181; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 3.6e-175;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 51 IRLGAFRLNLSNLSGGSTLTQOLIKITFTSTSDOTISRKAOEAWLAIOLEOKATK 110
DB 104 IRLGAFRLNLSNLSGGSTLTQOLIKITFTSTSDOTISRKAOEAWLAIOLEOKATK 163
OY 111 QEILFYINKVYMSNGNGMGTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHE 170
DB 164 QEILFYINKVYMSNGNGMGTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHE 223
OY 171 AAODRRNLVLESMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINOV 230
DB 224 AAODRRNLVLESMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINOV 283
OY 231 E 231
DB 284 E 284
```

```

RESULT 20
O54948 PRELIMINARY; PRT: 719 AA.
AC O54948;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DF 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=670;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain."
RL EMBL J. 11:3831-3836(1992).
DR EMBL; X67867; CAA48067.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
FT NON_TER 719
SQ SEQUENCE 719 AA; 79611 MW; 1DBFDDA93704930B CRC64;
```

Query Match 27.2%; Score 181; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 3.6e-175;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 51 IRLGAFRLNLSNLSGGSTLTQOLIKITFTSTSDOTISRKAOEAWLAIOLEOKATK 110
DB 104 IRLGAFRLNLSNLSGGSTLTQOLIKITFTSTSDOTISRKAOEAWLAIOLEOKATK 163
OY 111 QEILFYINKVYMSNGNGMGTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHE 170
DB 164 QEILFYINKVYMSNGNGMGTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHE 223
OY 171 AAODRRNLVLESMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINOV 230
DB 224 AAODRRNLVLESMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINOV 283
OY 231 E 231
DB 284 E 284
```

```

RESULT 21
O54949 PRELIMINARY; PRT: 719 AA.
AC O54949;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DF 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56742;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
```


RT South Africa and Spain.";
RL EMBL J. 11:3831-3836(1992).
DR EMBL; X67869; CAA48069.1; -;
DR InterPro; IPR001264; Transglycosyl.
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF00912; Transglycosyl; 1.
DR Pfam; PF00905; Transpeptidase; 1.
DR ProDom; PD001895; Transglycosyl; 1.
FT NON_TER 719 719
SQ SEQUENCE 719 AA; 79750 MW; 4B9B6C6ECB861CD CRC64;

Query Match 27.2%; Score 181; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 3.6e-175;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 51 IRLGAFRLNLSOGGTTLQOLIKLTFSTSDQTSRKAEAWLAIOLEKATK 110
DB 104 IRLGAFRLNLSOGGTTLQOLIKLTFSTSDQTSRKAEAWLAIOLEKATK 163
OY 111 QELITYYINKVYNSNGNYGQTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPE 170
DB 164 QELITYYINKVYNSNGNYGQTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPE 223
OY 171 AADRRNLVLEKNGKNGYISAEYEKAVNPTIDGLOSLKASANNYPAYMDNYKEVINY 230
DB 224 AADRRNLVLEKNGKNGYISAEYEKAVNPTIDGLOSLKASANNYPAYMDNYKEVINY 283
OY 231 E 231
DB 284 E 284

RESULT 22
O9EFO2 PRELIMINARY; PRT; 163 AA.
AC O9EFO2:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-950423; AND 960036;
RA Overweg K., Bogaert D., Sluifster M., de Groot R., Hermans P.W.M.;
RT "Molecular characterization of Streptococcus pneumoniae penicillin-
resistance in the Netherlands.";
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ403978; CAC20960.1; -;
DR EMBL; AJ403975; CAC20957.1; -;
FT NON_TER 163 163
FT NON_TER 163 163
SQ SEQUENCE 163 AA; 18477 MW; 99C44911EB96E5D3 CRC64;

Query Match 24.5%; Score 163; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 2e-157;
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 213 SNYPAYMDNYLKEVINQVEETGYNLLTGMVYTNVDOEAKHLMIDYNTDEYVYPPD 272
DB 1 SNYPAYMDNYLKEVINQVEETGYNLLTGMVYTNVDOEAKHLMIDYNTDEYVYPPD 60
OY 273 ELQVASTIVDSNGKVIAGLARGHSSNVSFGINQAVETNRDWSGTMKPTTDYAPALEYG 332
DB 61 ELQVASTIVDSNGKVIAGLARGHSSNVSFGINQAVETNRDWSGTMKPTTDYAPALEYG 120
OY 333 VYDSTATIVHDEPYNPGINTPYNMDRGYFGNITLQYALQOS 375

DB 121 VYDSTATIVHDEPYNPGINTPYNMDRGYFGNITLQYALQOS 163

RESULT 23
O9RET8 PRELIMINARY; PRT; 719 AA.
AC O9RET8:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SP 1465;
RC MEDLINE-21432820; PubMed-11549185;
RA Ferion A., Berche P.;
RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
penicillin-resistant clinical isolates of Streptococcus pneumoniae
serotype 23F from the nasopharyngeal flora of children.";
RL J. Med. Microbiol. 50:828-832(2001).
DR EMBL; AF210747; AAF17257.1; -;
DR InterPro; IPR001264; Transglycosyl.
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF00912; Transglycosyl; 1.
DR Pfam; PF00905; Transpeptidase; 1.
DR ProDom; PD001895; Transglycosyl; 1.
SQ SEQUENCE 719 AA; 79662 MW; A7B96EA15434A805 CRC64;

Query Match 21.3%; Score 142; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 1.9e-135;
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KIKDNKNQLIADGSRERRVNAQANDIPTDLKAIIVSIEDHREFDNGIDPTRLGAFLRN 60
DB 54 KIKDNKNQLIADGSRERRVNAQANDIPTDLKAIIVSIEDHREFDNGIDPTRLGAFLRN 113
OY 61 LQSNLSLOGGSTLTQOLIKLTFSTSDQTSRKAEAWLAIOLEKATKQELITYYINK 120
DB 114 LQSNLSLOGGSTLTQOLIKLTFSTSDQTSRKAEAWLAIOLEKATKQELITYYINK 173
OY 121 VYMSNGNYGQTAQNYGKDL 142
DB 174 VYMSNGNYGQTAQNYGKDL 195

RESULT 24
O9EW43 PRELIMINARY; PRT; 163 AA.
AC O9EW43:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-950421;
RA Overweg K., Bogaert D., Sluifster M., de Groot R., Hermans P.W.M.;
RT "Molecular characterization of Streptococcus pneumoniae penicillin-
resistance in the Netherlands.";
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ403977; CAC20959.1; -;

QY 560 VAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTSSSS 619
 DB 294 VAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTSSSS 353
 QY 620 SSDSSTSSSSTPSTNNSTTNPNNNTQOQSNTPDQONONPOPA 664
 DB 354 SSDSSTSSSSTPSTNNSTTNPNNNTQOQSNTPDQONONPOPA 398

RESULT 28

Q9R7M7 PRELIMINARY: PRT: 398 AA.

AC Q9R7M7
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=7/K133;
 RX MEDLINE=98409715; PubMed=9736547;
 RA Asahi Y., Ubukata K.;
 RT "Association of a Thr-371 substitution in a conserved amino acid motif of penicillin-binding protein 1A with penicillin resistance of Streptococcus pneumoniae."
 RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
 DR EMBL; AB006872; BAA32068.1;
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00905; Transpeptidase; 1.
 FT NON_TER 1
 FT NON_TER 398
 SQ SEQUENCE 398 AA; 43638 MW; 551ADB567C3F79BE CRC64;

Query Match 15.8%; Score 105; DB 2; Length 398;
 Best Local Similarity 100.0%; Pred. No. 5.3e-98;
 Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 VAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTSSSS 619
 DB 294 VAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTSSSS 353
 QY 620 SSDSSTSSSSTPSTNNSTTNPNNNTQOQSNTPDQONONPOPA 664
 DB 354 SSDSSTSSSSTPSTNNSTTNPNNNTQOQSNTPDQONONPOPA 398

RESULT 29

Q87105 PRELIMINARY: PRT: 398 AA.

ID Q87105
 AC Q87105;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9/217;
 RX MEDLINE=98409715; PubMed=9736547;
 RA Asahi Y., Ubukata K.;
 RT "Association of a Thr-371 substitution in a conserved amino acid motif of penicillin-binding protein 1A with penicillin resistance of Streptococcus pneumoniae."
 FT streptococcus pneumoniae.";

RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
 DR EMBL; AB006873; BAA32069.1;
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00905; Transpeptidase; 1.
 FT NON_TER 1
 FT NON_TER 398
 SQ SEQUENCE 398 AA; 43727 MW; 3AA038C8B5582373 CRC64;

Query Match 15.8%; Score 105; DB 2; Length 398;

Best Local Similarity 100.0%; Pred. No. 5.3e-98; Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 VAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTSSSS 619
 DB 294 VAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTSSSS 353
 QY 620 SSDSSTSSSSTPSTNNSTTNPNNNTQOQSNTPDQONONPOPA 664
 DB 354 SSDSSTSSSSTPSTNNSTTNPNNNTQOQSNTPDQONONPOPA 398

RESULT 30

Q54918 PRELIMINARY: PRT: 189 AA.

ID Q54918
 AC Q54918;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE INTERNAL REGION OF THE PENICILLIN-BINDING PROTEIN 1A GENE (FRAGMENT).
 GN PBP 1A.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=85983 (PENICILLIN-RESISTANT STRAIN);
 RX MEDLINE=9401131; PubMed=8406829;
 RA Kell C.M., Jordens Z., Daniels M., Coffey T.J., Bates J., Paul J., Glicks C., Spratt B.G.;
 RT "Molecular epidemiology of penicillin-resistant pneumococci isolated in Nairobi, Kenya."
 RL Infect. Immun. 61:4382-4391(1993).
 DR EMBL; Z21800; CAA79864.1;
 FT NON_TER 1
 FT NON_TER 189
 SQ SEQUENCE 189 AA; 20740 MW; ADDDBCC22322AA06D CRC64;

Query Match 14.7%; Score 98; DB 2; Length 189;
 Best Local Similarity 100.0%; Pred. No. 3.6e-91;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 557 GLTVAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTSS 616
 DB 92 GLTVAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTSS 151
 QY 617 SSSSSDSTSSSSTPSTNNSTTNPNNNTQOQSNTP 654
 DB 152 SSSSSDSTSSSSTPSTNNSTTNPNNNTQOQSNTP 189

RESULT 31

Q52737 PRELIMINARY: PRT: 310 AA.

ID Q52737
 AC Q52737;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
 GN PONA.

OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=65654;
RX MEDLINE=98287565; PubMed=9624469;
RA Smith A.M., Klugman K.P.;
RT "Alterations in PBP 1A essential for high-level penicillin resistance
in Streptococcus pneumoniae.";
RL Antimicrob. Agents Chemother. 42:1329-1333(1998).
DR EMBL: AF046231; AAC24696.1; -
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00905; Transpeptidase; 1.
FT NON_TER 1
FT NON_TER 310 310
SQ SEQUENCE 310 AA; 34419 MW; 7C683053109F6DA CRC64;

Query Match 14.3%; Score 95; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 6.5e-68;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 LMDIYNDEYVAYPPDELOVASTIVDSNGKVIQOLGARHSSNVSGINQAVETNRDVG 316
Db 1 LMDIYNDEYVAYPPDELOVASTIVDSNGKVIQOLGARHSSNVSGINQAVETNRDVG 60
QY 317 STMKPTIDYAPALEXGYVDSTATIVHDEPYNPGT 351
Db 61 STMKPTIDYAPALEXGYVDSTATIVHDEPYNPGT 95

RESULT 32
Q9R7M5 PRELIMINARY; PRT; 398 AA.
AC Q9R7M5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=#24/TU25; #26/TU29;
RX MEDLINE=98409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a Thr-371 substitution in a conserved amino acid motif
of penicillin-binding protein 1A with penicillin resistance of
Streptococcus pneumoniae.";
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
DR EMBL: AB006879; BAA32075.1; -
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00905; Transpeptidase; 1.
FT NON_TER 1
FT NON_TER 398 398
SQ SEQUENCE 398 AA; 43654 MW; D1E181574F2B58B8 CRC64;

Query Match 12.8%; Score 85; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 1.2e-77;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 VAYPDELOVASTIVDSNGKVIQOLGARHSSNVSGINQAVETNRDVGSTMKPTIDYA 326
Db 1 VAYPDELOVASTIVDSNGKVIQOLGARHSSNVSGINQAVETNRDVGSTMKPTIDYA 60
QY 327 PALEXGYVDSTATIVHDEPYNPGT 351

Db 61 PALEXGYVDSTATIVHDEPYNPGT 85

RESULT 33
O87108 PRELIMINARY; PRT; 398 AA.
AC O87108;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=#23/HSB21;
RX MEDLINE=98409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a Thr-371 substitution in a conserved amino acid motif
of penicillin-binding protein 1A with penicillin resistance of
Streptococcus pneumoniae.";
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
DR EMBL: AB006878; BAA32074.1; -
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00905; Transpeptidase; 1.
FT NON_TER 1
FT NON_TER 398 398
SQ SEQUENCE 398 AA; 43675 MW; F513E2C548DC28AA CRC64;

Query Match 12.0%; Score 80; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 1.5e-72;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 407 HYSNMISSNTESDKKYGSSSEKMAAAYAFANGSTYKPMYIHKVVSDDGSEKFSNVG 466
Db 141 HYSNMISSNTESDKKYGSSSEKMAAAYAFANGSTYKPMYIHKVVSDDGSEKFSNVG 200
QY 467 TRAMKETAYMTDMKKTVL 486
Db 201 TRAMKETAYMTDMKKTVL 220

RESULT 34
O54950 PRELIMINARY; PRT; 719 AA.
AC O54950;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2039;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
clones of penicillin-resistant Streptococcus pneumoniae isolated in
South Africa and Spain.";
RL EMBL J. 11:3831-3836(1992).
DR EMBL: X67870; CAA48070.1; -
DR InterPro: IPR001264; Transglycosyl.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.

DR ProDom: PD001895; Transglycosyl; 1.
FT NON_TER 719 719
SQ SEQUENCE 719 AA; 79874 MW; FACBPAB303D865D4 CRC64;

Query Match 11.9%; Score 79; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 2.8e-71;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 NSLOGSSTLTQOLIKTFYFSTSTSDQFISRKQEAWLAIQLEQKATKQELTYINKYVM 123
DB 117 NSLOGSSTLTQOLIKTFYFSTSDQFISRKQEAWLAIQLEQKATKQELTYINKYVM 176

QY 124 SNGNYGQTAQNYGKDL 142
DB 177 SNGNYGQTAQNYGKDL 195

RESULT 35
087107
AC 087107 PRELIMINARY; PRT: 398 AA.

DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=422/HA5.
RX MEDLINE=98409715; PubMed-9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a Thr-371 substitution in a conserved amino acid motif
of penicillin-binding protein 1A with penicillin resistance of
RT Streptococcus pneumoniae."
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
DR EMBL; AB006877; BAA32073.1; -;
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF00905; Transpeptidase; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 398 AA; 43487 MW; CDF82616C5159991 CRC64;

Query Match 11.4%; Score 76; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 1.8e-68;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 589 RGEFFKNGASTWNSPAPQPPSTESSSSSDSTSSSSTPTNNSTTNPNNNTQ 648
DB 323 RGEFFKNGASTWNSPAPQPPSTESSSSSDSTSSSSTPTNNSTTNPNNNTQ 382

QY 649 OSNTTPDOONONPOPA 664
DB 383 OSNTTPDOONONPOPA 398

RESULT 36
09EM42
ID 09EM42 PRELIMINARY; PRT: 163 AA.

AC 09EM42;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.

OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=950473.
RA Overweg K., Bogaert D., Sluifster M., de Groot R., Hermans P.W.M.;
RT "Molecular characterization of Streptococcus pneumoniae penicillin-
RT resistance in the Netherlands."
RL Substituted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ403979; CAC20961.1; -;
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 163 AA; 18350 MW; 02B96CAE2DB26ACF CRC64;

Query Match 9.0%; Score 60; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 1.6e-52;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 SNYPAYMDNYLKEVINQVEEETGYNLLTGMQDYYTNVDQEAQRHMDIYNTDEYVAPDD 272
DB 1 SNYPAYMDNYLKEVINQVEEETGYNLLTGMQDYYTNVDQEAQRHMDIYNTDEYVAPDD 60

RESULT 37
09R7M8
ID 09R7M8 PRELIMINARY; PRT: 398 AA.

AC 09R7M8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=43/RH2, #6/KU126, #15/Z12, #16/Z13;
RX MEDLINE=98409715; PubMed-9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a Thr-371 substitution in a conserved amino acid motif
of penicillin-binding protein 1A with penicillin resistance of
RT Streptococcus pneumoniae."
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
DR EMBL; AB006870; BAA32066.1; -;
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF00905; Transpeptidase; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 398 AA; 43583 MW; 0BD9B32C6B2D577 CRC64;

Query Match 9.0%; Score 60; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 3.6e-52;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 SPAPQPPSTESSSSSDSTSSSSTPTNNSTTNPNNNTQOSNTTPDOONONPOPA 664
DB 339 SPAPQPPSTESSSSSDSTSSSSTPTNNSTTNPNNNTQOSNTTPDOONONPOPA 398

RESULT 38
09F2G4
ID 09F2G4 PRELIMINARY; PRT: 727 AA.

AC 09F2G4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PENIA.
OS Streptococcus mitis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Streptococcus.
OX NCBI_TaxID=28037;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=209;
RX MEDLINE=21301162; PubMed=11408226;
RA Amoroso A., Demares D., Mollerach M., Gutkind G., Coyette J.;
RT "All detectable high-molecular-mass penicillin-binding proteins are
modified in a high-level beta-lactam-resistant clinical isolate of
Streptococcus mitis";
RL Antimicrob. Agents Chemother. 45:2075-2081(2001).
DR EMBL; AJ295856; CAC08466.1; -;
DR InterPro; IPR001264; Transglycosyl.
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF009012; Transglycosyl; 1.
DR Pfam; PF009005; Transpeptidase; 1.
DR Prodom; PD001895; Transglycosyl; 1.
FT NON_TER 727 727
SQ SEQUENCE 727 AA; 80506 MW; AFF56454A4E30129 CRC64;

Query Match 8.1%; Score 54; DB 2; Length 727;
Best Local Similarity 100.0%; Pred. No. 8.2e-46;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 285 NGKVTAGLGAHROSNSVSGINQAVETNRDMSGTMKPTIDYAPALEGYVOSTA 338
DB 339 NGKVTAGLGAHROSNSVSGINQAVETNRDMSGTMKPTIDYAPALEGYVOSTA 392

RESULT 39
087104
ID 087104 PRELIMINARY; PRT; 398 AA.

AC 087104;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=#5/H31, #8/22, #11/220, #12/221;
RX MEDLINE=96409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a Thr-371 substitution in a conserved amino acid motif
of penicillin-binding protein 1A with penicillin resistance of
Streptococcus pneumoniae";
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
DR EMBL; AB006871; BAA32067.1; -;
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF009005; Transpeptidase; 1.
FT NON_TER 1 1
SQ SEQUENCE 398 AA; 43536 MW; 122DOB05F8EF117C CRC64;

Query Match 8.0%; Score 53; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 4.9e-45;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 569 RGEFVFNGARSTWNSPAPQOPSTESSSSSDSTQSSSTPSTNNSTTT 641
DB 323 RGEFVFNGARSTWNSPAPQOPSTESSSSSDSTQSSSTPSTNNSTTT 375

RESULT 40
O9EM44 PRELIMINARY; PRT; 163 AA.
ID O9EM44
AC O9EM44;

DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=950710;
RA Overweg K., Bogaert D., Sluifder M., de Groot R., Hermans P.W.M.;
RT "Molecular characterization of Streptococcus pneumoniae penicillin-
resistance in the Netherlands";
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ403976; CAC20958.1; -;
FT NON_TER 1 1
SQ SEQUENCE 163 AA; 18358 MW; 3D50DD342D9A0900 CRC64;

Query Match 7.5%; Score 50; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 2.4e-42;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 213 SNYPAMDNYLKEVINQVEEETGYNLLTGMQDYTVNVDQAKHMDIYN 262
DB 1 SNYPAMDNYLKEVINQVEEETGYNLLTGMQDYTVNVDQAKHMDIYN 50

RESULT 41
O9EM41
ID O9EM41 PRELIMINARY; PRT; 163 AA.

AC O9EM41;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=960027;
RA Overweg K., Bogaert D., Sluifder M., de Groot R., Hermans P.W.M.;
RT "Molecular characterization of Streptococcus pneumoniae penicillin-
resistance in the Netherlands";
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ403980; CAC20962.1; -;
FT NON_TER 1 1
SQ SEQUENCE 163 AA; 18464 MW; 83007AD7ACB00B31 CRC64;

Query Match 6.2%; Score 41; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 3.6e-33;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 253 AOKHLMIDYNTDEYAVYPPDELOVASTIVDSNGKVIADLG 293
DB 41 AOKHLMIDYNTDEYAVYPPDELOVASTIVDSNGKVIADLG 81

RESULT 42
O93N72
ID O93N72 PRELIMINARY; PRT; 310 AA.
AC O93N72;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=45904;
 RA MEDLINE=21393681; PubMed=11502545;
 RX Smith A.M., Botha R.F., Koorhof H.J., Klugman K.P.;
 RT "Emergence of a Pneumococcal Clone with Cephalosporin Resistance and
 Penicillin Susceptibility.";
 RL Antimicrob. Agents Chemother. 45:2648-2650(2001).
 DR EMBL; AF387163; AAK93962.1; -.
 FT NON_TER 1 1
 SQ SEQUENCE 310 AA; 34249 MW; 9756459BCFA2EF CRC64;

Query Match 6.2%; Score 41; DB 2; Length 310;
 Best Local Similarity 100.0%; Pred. No. 6.5e-33;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 257 LMDIYNTDEVAYPPDELOVASTIVDSNGKYIAQIGARHQ 297
 DB 1 LMDIYNTDEVAYPPDELOVASTIVDSNGKYIAQIGARHQ 41

RESULT 43
 ID 070037 PRELIMINARY; PRT; 527 AA.
 AC 070037;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
 GN PBPIA.
 OS Streptococcus mitis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=28037;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B6;
 RA Hakenbeck R., Koenig A., Kern I., van der Linden M., Keck W.,
 RA Billot-Klein D., Legrand R., Schoot B., Gutmann L.;
 RT "Acquisition of five high-Mr penicillin-binding protein variants
 during transfer of high-level beta-lactam resistance from
 Streptococcus mitis to Streptococcus pneumoniae.";
 RT J. Bacteriol. 180:1831-1840(1998).
 DR EMBL; AJ002290; CAA05301.1; -.
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 FT NON_TER 1 1
 SQ SEQUENCE 527 AA; 58859 MW; 8952B38713827021 CRC64;

Query Match 5.7%; Score 38; DB 2; Length 527;
 Best Local Similarity 100.0%; Pred. No. 1.2e-29;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 NSLOGGSLTQOLIKITFTSTSDQITSRKAQEAFLA 101
 DB 67 NSLOGGSLTQOLIKITFTSTSDQITSRKAQEAFLA 104

RESULT 44
 O9F2G7

ID O9F2G7 PRELIMINARY; PRT; 728 AA.
 AC O9F2G7;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
 GN PENIA.
 OS Streptococcus mitis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=28037;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=127R;
 RA MEDLINE=21301162; PubMed=11408226;
 RX Amoroso A., Demares D., Moellerach M., Gutkind G., Coyette J.;
 RT "All detectable high-molecular-mass penicillin-binding proteins are
 modified in a high-level beta-lactam-resistant clinical isolate of
 Streptococcus mitis.";
 RT Antimicrob. Agents Chemother. 45:2075-2081(2001).
 DR EMBL; AJ295852; CAC08463.1; -.
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 FT NON_TER 728 728
 SQ SEQUENCE 728 AA; 80854 MW; 37E110286600987 CRC64;

Query Match 4.8%; Score 32; DB 2; Length 728;
 Best Local Similarity 100.0%; Pred. No. 2.1e-23;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 103 QLEQKATKOEILTYINKYMSNGYGMOTAA 134
 DB 157 QLEQKATKOEILTYINKYMSNGYGMOTAA 188

RESULT 45
 ID O9EW45 PRELIMINARY; PRT; 163 AA.
 AC O9EW45;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
 DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=950225;
 RA Overweg K., Bogaert D., Sluifker M., de Groot R., Hermans P.W.M.;
 RT "Molecular characterization of Streptococcus pneumoniae penicillin-
 resistance in the Netherlands.";
 RT Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF403974; CAC20956.1; -.
 FT NON_TER 1 1
 SQ SEQUENCE 163 AA; 18357 MW; F6DED7F42D80D3CB CRC64;

Query Match 4.5%; Score 30; DB 2; Length 163;
 Best Local Similarity 100.0%; Pred. No. 5.7e-22;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 233 ETGYNLLTGMVYTNVDOEAKHMDIYN 262
 DB 21 ETGYNLLTGMVYTNVDOEAKHMDIYN 50

RESULT 46
087106 PRELIMINARY; PRT; 400 AA.
ID 087106;
AC 087106;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-#14/242;
RX MEDLINE=98409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a Thr-371 substitution in a conserved amino acid motif
of penicillin-binding protein 1A with penicillin resistance of
Streptococcus pneumoniae."
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
DR EMBL; AB006875; BAA32071.1; -
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00905; Transpeptidase; 1.
FT NON_TER 1 1
FT SEQUENCE 400 AA; 43773 MW; 0B8965011453D226 CRC64;
SQ
Query Match 4.1%; Score 27; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 1,5e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 267 VAYPDELQVASTIVDSNGKVIADLG 293
Db 1 VAYPDELQVASTIVDSNGKVIADLG 27
OY 267 VAYPDELQVASTIVDSNGKVIADLG 293
Db 1 VAYPDELQVASTIVDSNGKVIADLG 27
RESULT 47
099YL1 PRELIMINARY; PRT; 721 AA.
ID 099YL1;
AC 099YL1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PURATIVE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA OR SPY1649.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., Moshan W.M., Ajdic D.J., Savic G., Lyon K.,
RA Primeaux C., Szatse S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL; AE006596; AAK34416.1; -
DR InterPro: IPR001264; Transglycosyl.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
KW Complete proteome.
SQ SEQUENCE 721 AA; 80207 MW; 7E9350705EF4C6D0 CRC64;

Query Match 3.5%; Score 23; DB 16; Length 721;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 419 SPKKYGASSEKMAAAYAFANGG 441
Db 475 SDKKYGASSEKMAAAYAFANGG 497
OY 419 SPKKYGASSEKMAAAYAFANGG 441
Db 475 SDKKYGASSEKMAAAYAFANGG 497
RESULT 48
09EW40 PRELIMINARY; PRT; 163 AA.
ID 09EW40;
AC 09EW40;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-960035;
RA Overweg K., Bogaert D., Sluifder M., de Groot R., Hermans P.W.M.;
RT "Molecular characterization of Streptococcus pneumoniae penicillin-
resistance in the Netherlands."
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ403981; CAC20963.1; -
FT NON_TER 1 1
FT NON_TER 1 1
FT SEQUENCE 163 AA; 18347 MW; BC800A043C6E1FCF CRC64;
SQ
Query Match 3.3%; Score 22; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 8e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 241 TGMVYTNVDQEAOKHMDIYN 262
Db 29 TGMVYTNVDQEAOKHMDIYN 50
OY 241 TGMVYTNVDQEAOKHMDIYN 262
Db 29 TGMVYTNVDQEAOKHMDIYN 50
RESULT 49
054919 PRELIMINARY; PRT; 139 AA.
ID 054919;
AC 054919;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE INTERNAL REGION OF THE PENICILLIN-BINDING PROTEIN 1A GENE
(FRAGMENT).
GN PBP 1A.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-100511 (PENICILLIN-RESISTANT STRAIN);
RX MEDLINE=94011331; PubMed=8406829;
RA Kell C.M., Jordens Z., Daniels M., Coffey T.J., Bates J., Paul J.,
RA Gilks C., Spratt B.G.;
RT "Molecular epidemiology of penicillin-resistant pneumococci isolated
in Nairobi, Kenya."
RL Infect. Immun. 61:4382-4391(1993).
DR EMBL; Z21809; CAA79874.1; -
FT NON_TER 1 1
FT NON_TER 1 1
FT SEQUENCE 139 AA; 15250 MW; 40ED8ED03593D446 CRC64;

Query Match 3.0%; Score 20; DB 2; Length 139;
 Best Local Similarity 100.0%; Pred. No. 7.6e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 533 GYTRKYSMAVMTGYSNRLTP 552
 ||||||||||||||||
 Db 16 GYTRKYSMAVMTGYSNRLTP 35

RESULT 50

052744 PRELIMINARY; PRT; 310 AA.
 ID 052744
 AC 052744;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
 GN PONA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=63509, M11;
 RX MEDLINE=98287565; PubMed=9624469;
 RA Smith A.M., Klugman K.P.;
 RT "Alterations in PBP 1A essential for high-level penicillin resistance
 in Streptococcus pneumoniae."
 RL Antimicrob. Agents Chemother. 42:1329-1333(1998).
 DR EMBL; AF046238; AAC24703.1; -;
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF00905; Transpeptidase; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 310 AA; 34354 MW; F8FC8AC00E7BF5FF CRC64;

Query Match 3.0%; Score 20; DB 2; Length 310;
 Best Local Similarity 100.0%; Pred. No. 1.6e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 MKPITDYAPALEGYVDSTA 338
 ||||||||||||||||
 Db 63 MKPITDYAPALEGYVDSTA 82

Search completed: June 13, 2002, 08:50:23
 Job time: 238 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:42:44 ; Search time 14.67 seconds
(without alignments)
1108.892 Million cell updates/sec

Title: US-08-961-083-2
Perfect score: 666
Sequence: 1 KIYDNKNQLIADLGSERRRN.....TOOSNTTPOOONPOPAQP 666

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 10

Total number of hits satisfying chosen parameters: 8

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 1000 summaries

Database :

Issued_Patents_AA*
1: /cgn2-6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2-6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2-6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2-6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2-6/ptodata/2/1aa/PTUS.COMB.pep:*
6: /cgn2-6/ptodata/2/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	666	100.0	666	US-08-961-083-2	Sequence 2, Appli
2	330	49.5	682	US-08-481-435-6	Sequence 6, Appli
3	245	36.8	320	US-08-245-511-4	Sequence 4, Appli
4	245	36.8	320	US-08-993A-4	Sequence 4, Appli
5	77	11.6	77	US-08-245-511-24	Sequence 24, Appli
6	77	11.6	77	US-08-600-993A-24	Sequence 24, Appli
7	10	1.5	110	US-08-961-083-102	Sequence 102, App
8	10	1.5	1704	US-08-485-355B-40	Sequence 40, Appli

ALIGNMENTS

RESULT 1
US-08-961-083-2
Sequence 2, Application US/08961083
Patent No. 6159469

GENERAL INFORMATION:

APPLICANT: Chol et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB34072
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-2

Query Match 100.0%, Score 666; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KIYDNKNQLIADLGSERRRNAOANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLENN	60
DB	1	KIYDNKNQLIADLGSERRRNAOANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLENN	60
QY	61	IQSNLSGGSTLTQOLIKITFYSTSDOTISRKQAEMLAIQLEKAKKQELTYIYNNK	120
DB	61	IQSNLSGGSTLTQOLIKITFYSTSDOTISRKQAEMLAIQLEKAKKQELTYIYNNK	120
QY	121	VYMSNGNYGMOTAAQNYVYKGLNNLSLPOLALLAGPQAPNOYDPSHPEAODRNNLY	180
DB	121	VYMSNGNYGMOTAAQNYVYKGLNNLSLPOLALLAGPQAPNOYDPSHPEAODRNNLY	180
QY	181	SEMKNQYISADQYERAVNPTITDGLQSLKSASNPAYMDNYLKEYINQVEERTGYNLT	240
DB	181	SEMKNQYISADQYERAVNPTITDGLQSLKSASNPAYMDNYLKEYINQVEERTGYNLT	240
QY	241	TGMQYVYNNDOAKHMDIYNTDEVAVPDELOVASTIVVSNKVIYAOIGARHOSN	300
DB	241	TGMQYVYNNDOAKHMDIYNTDEVAVPDELOVASTIVVSNKVIYAOIGARHOSN	300
QY	301	VSEGINQAVETNRDNGSTPMKPTIDYAPALEYGVYOSTATIVHDEPNYPGTNPNMNR	360
DB	301	VSEGINQAVETNRDNGSTPMKPTIDYAPALEYGVYOSTATIVHDEPNYPGTNPNMNR	360
QY	361	GYFGNTTQYALQOSRNPVAVETLNKVGINRAKTEFLNGIDYPSITHYSNAISSNTTESD	420
DB	361	GYFGNTTQYALQOSRNPVAVETLNKVGINRAKTEFLNGIDYPSITHYSNAISSNTTESD	420
QY	421	KRYGASSEKMAAAYAFAGNGTYRPMYTHKVFSDGSEKESNNGTRAMKETTAAMMD	480
DB	421	KRYGASSEKMAAAYAFAGNGTYRPMYTHKVFSDGSEKESNNGTRAMKETTAAMMD	480
QY	481	NKTYVLTGTRNNAYLAMPQAGKTGTSNYTDEELENHKTQFVAPDELFGYTRKYSN	540
DB	481	NKTYVLTGTRNNAYLAMPQAGKTGTSNYTDEELENHKTQFVAPDELFGYTRKYSN	540
QY	541	AVMTGYSNRLPLVNGILVAAKVVYSMTYLSGSPNEDMNIPEGLYRNGSEFVRNGAR	600
DB	541	AVMTGYSNRLPLVNGILVAAKVVYSMTYLSGSPNEDMNIPEGLYRNGSEFVRNGAR	600

Best Local Similarity 100.0%; Pred. No. 1.3e-241;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LKQQLIKLFFSTSTSDQITSRKAOEAWLAIOLEKATKOEILTYINKVYNSNGNYGMQ 131
Db 76 LKQQLIKLFFSTSTSDQITSRKAOEAWLAIOLEKATKOEILTYINKVYNSNGNYGMQ 135
OY 132 TAAQNTYGGDLNLSLPQALLAGMPQAPNOYDPSHPPEAODRRNLVSEKMGYISA 191
Db 136 TAAQNTYGGDLNLSLPQALLAGMPQAPNOYDPSHPPEAODRRNLVSEKMGYISA 195
OY 192 EYERKAVNPTIDGLOSLSKSNYPAYMDNYLKEVINQVEEETGYNLTTGMDVYTNVQ 251
Db 196 EYERKAVNPTIDGLOSLSKSNYPAYMDNYLKEVINQVEEETGYNLTTGMDVYTNVQ 255
OY 252 EAQKHLMDIYNTEDEYAYVPPDELOVASTIVDVSNGKVIQAOLGARHSSNVSGINQAVET 311
Db 256 EAQKHLMDIYNTEDEYAYVPPDELOVASTIVDVSNGKVIQAOLGARHSSNVSGINQAVET 315
OY 312 NRDMG 316
Db 316 NRDMG 320

RESULT 4

US-08-600-993A-4
; Sequence 4, Application US/08600993A

; Patent No. 5981229

; GENERAL INFORMATION:

; APPLICANT: Masure, H Robert

; APPLICANT: Pearce, Barbara J

; APPLICANT: Tuomanen, Elaine

; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND

; NUMBER OF SEQUENCES: 59

; CORRESPONDENCE ADDRESS:

; ADDRESS: Klauber & Jackson

; STREET: 411 Hackensack Avenue

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/600.993A

; FILING DATE: 1-MAR-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/245.511

; FILING DATE: 18-MAY-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/116.541

; FILING DATE: 01-SEP-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.

; REGISTRATION NUMBER: 26,742

; REFERENCE/DOCKET NUMBER: 600-1-069 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201 487-5800

; TELEFAX: 201 343-1684

; TELEFAX: 133521

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 320 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-600-993A-4

Query Match

36.8%; Score 245; DB 2; Length 320;

Best Local Similarity 100.0%; Pred. No. 1.3e-241;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LKQQLIKLFFSTSTSDQITSRKAOEAWLAIOLEKATKOEILTYINKVYNSNGNYGMQ 131
Db 76 LKQQLIKLFFSTSTSDQITSRKAOEAWLAIOLEKATKOEILTYINKVYNSNGNYGMQ 135
OY 132 TAAQNTYGGDLNLSLPQALLAGMPQAPNOYDPSHPPEAODRRNLVSEKMGYISA 191
Db 136 TAAQNTYGGDLNLSLPQALLAGMPQAPNOYDPSHPPEAODRRNLVSEKMGYISA 195
OY 192 EYERKAVNPTIDGLOSLSKSNYPAYMDNYLKEVINQVEEETGYNLTTGMDVYTNVQ 251
Db 196 EYERKAVNPTIDGLOSLSKSNYPAYMDNYLKEVINQVEEETGYNLTTGMDVYTNVQ 255
OY 252 EAQKHLMDIYNTEDEYAYVPPDELOVASTIVDVSNGKVIQAOLGARHSSNVSGINQAVET 311
Db 256 EAQKHLMDIYNTEDEYAYVPPDELOVASTIVDVSNGKVIQAOLGARHSSNVSGINQAVET 315
OY 312 NRDMG 316
Db 316 NRDMG 320

RESULT 5

US-08-245-511-24
; Sequence 24, Application US/08245511

; Patent No. 5928900

; GENERAL INFORMATION:

; APPLICANT: Masure, H Robert

; APPLICANT: Tuomanen, Elaine

; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND

; NUMBER OF SEQUENCES: 58

; CORRESPONDENCE ADDRESS:

; ADDRESS: Klauber & Jackson

; STREET: 411 Hackensack Avenue

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/245.511

; FILING DATE: 18-MAY-1994

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/116.541

; FILING DATE: 01-SEP-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.

; REGISTRATION NUMBER: 26,742

; REFERENCE/DOCKET NUMBER: 600-1-069 CIP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201 487-5800

; TELEFAX: 201 343-1684

; TELEFAX: 133521

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 77 amino acids

; TYPE: amino acid

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; US-08-245-511-24

FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
STRAIN: R6
IMMEDIATE SOURCE:
CLONE: SPR042
US-08-245-511-24

Query Match 11.6%; Score 77; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 8.1e-71;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 TTGMDVYTNVDOEAQKHLMDIYNTDEVAVYPDELQVASTIVDVNSGKVIAGLGARHSS 299
DB 1 TTGMDVYTNVDOEAQKHLMDIYNTDEVAVYPDELQVASTIVDVNSGKVIAGLGARHSS 60

QY 300 NVSFGINQAVETNRDMG 316
DB 61 NVSFGINQAVETNRDMG 77

RESULT 6

US-08-600-993A-24
Sequence 24, Application US/08600993A
Patent No. 5981229

GENERAL INFORMATION:

APPLICANT: Measure, H Robert
APPLICANT: Tuomanen, Elaine
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,993A
FILING DATE: 1-MAR-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,511
FILING DATE: 18-MAY-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/116,541
FILING DATE: 01-SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-069 US

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 77 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
STRAIN: R6
IMMEDIATE SOURCE:
CLONE: SPR042
US-08-600-993A-24

Query Match 11.6%; Score 77; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 8.1e-71;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 TTGMDVYTNVDOEAQKHLMDIYNTDEVAVYPDELQVASTIVDVNSGKVIAGLGARHSS 299
DB 1 TTGMDVYTNVDOEAQKHLMDIYNTDEVAVYPDELQVASTIVDVNSGKVIAGLGARHSS 60

QY 300 NVSFGINQAVETNRDMG 316
DB 61 NVSFGINQAVETNRDMG 77

RESULT 7

US-08-961-083-102
Sequence 102, Application US/08961083
Patent No. 6159469

GENERAL INFORMATION:

APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44M storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P2

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:

LENGTH: 110 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-961-083-102

Query Match 1.5%; Score 10; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
DB 18 SSSSSSDSST 27

RESULT 8
US-08-485-355B-40
; Sequence 40, Application US/08485355B
; Patent No. 6177075
; GENERAL INFORMATION:
; APPLICANT: Christlan, P. D., Gordon, K. H.J., Hanzlik, T. N.
; TITLE OF INVENTION: Insect Viruses and Their Uses in
; Protecting Plants
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,355B
; FILING DATE: 07-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,522
; FILING DATE: 12-MAY-1995
; APPLICATION NUMBER: US 08/089,372
; FILING DATE: 08-JUL-1993
; APPLICATION NUMBER: AU PL4081/92
; FILING DATE: 14-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-58631-2/RT/DSS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEEX: 910 277299
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1704 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 40:
; US-08-485-355B-40

Query Match 1.5%; Score 10; DB 4; Length 1704;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 617 SSSSSDSSTS 626
|||||
DB 1452 SSSSSDSSTS 1461

Search completed: June 13, 2002, 08:46:41
Job time: 237 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:44:44 ; Search time 113.7 Seconds
(without alignments)
2061.733 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 666

Sequence: 1 KIDYKNGQIADIGSERRVN.....TQSSNTTPDQGNQNPQAPQ 666

Scoring table:

GAPOP 60.0 , Gapext 60.0

Searched: 3502263 seqs, 351980561 residues

Word size : 10

Total number of hits satisfying chosen parameters: 88

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US087.COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US088.COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US089.COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US090.COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US091.COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US092.COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US093.COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US094.COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US095.COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US096.COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US097.COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US098.COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US099.COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US100.COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US101.COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US102.COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US103.COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US104.COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US105.COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US106.COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US107.COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US108.COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US109.COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US110.COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US111.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	666	100.0	666	19	US-09-536-784-2
2	666	100.0	666	21	US-09-765-271-2
3	666	100.0	666	21	US-09-765-272-2
4	565	84.8	719	19	US-09-583-110-3973
5	565	84.8	721	15	US-09-107-433-3705
6	330	49.5	719	1	PCT-US02-03987-13423
7	330	49.5	719	22	US-09-815-242-13423

8	330	49.5	719	24	US-10-072-851-13423	Sequence 13423, A
9	245	36.8	777	5	US-08-116-541-4	Sequence 4, App1
10	277	11.6	823	15	US-09-107-532-5667	Sequence 24, App1
11	18	2.7	823	15	US-09-107-532-5667	Sequence 5667, Ap
12	18	2.7	823	15	US-09-107-532A-5667	Sequence 5667, Ap
13	13	2.1	771	20	US-09-634-238-229	Sequence 229, App
14	14	2.1	771	20	US-09-634-238-262	Sequence 262, App
15	14	2.1	778	1	PCT-US02-03987-10728	Sequence 10728, A
16	14	2.1	778	22	US-09-815-242-10728	Sequence 10728, A
17	14	2.1	778	24	US-10-072-851-10728	Sequence 10728, A
18	14	2.1	789	15	US-09-134-000-4939	Sequence 4939, Ap
19	19	1.7	266	13	PCT-US97-14436-535	Sequence 535, App
20	11	1.7	266	13	US-08-911-503-535	Sequence 535, App
21	11	1.7	266	13	US-08-911-503A-535	Sequence 535, App
22	11	1.7	314	18	US-09-417-507-35782	Sequence 35782, A
23	11	1.7	820	19	US-09-583-110-4219	Sequence 4219, Ap
24	11	1.7	821	17	US-09-377-465-2	Sequence 2, App1
25	11	1.7	821	17	US-09-377-465A-2	Sequence 2, App1
26	11	1.7	834	15	US-09-107-433-4760	Sequence 4760, Ap
27	11	1.7	76	20	US-09-621-976-7706	Sequence 7706, Ap
28	10	1.5	76	22	US-09-834-366-24930	Sequence 24930, A
29	10	1.5	76	22	US-09-834-366-24930	Sequence 24930, A
30	10	1.5	76	26	US-09-147-499-7706	Sequence 7706, Ap
31	10	1.5	76	26	US-09-197-873-24929	Sequence 24929, A
32	10	1.5	76	26	US-09-197-873-24930	Sequence 24930, A
33	10	1.5	110	19	US-09-536-784-102	Sequence 102, App
34	10	1.5	110	21	US-09-765-271-102	Sequence 102, App
35	10	1.5	110	21	US-09-765-272-102	Sequence 102, App
36	10	1.5	119	20	US-09-614-150-34311	Sequence 34311, A
37	10	1.5	119	20	US-09-191-637-33895	Sequence 33895, A
38	10	1.5	130	20	US-09-621-976-5866	Sequence 5866, Ap
39	10	1.5	130	26	US-09-147-499-5866	Sequence 5866, Ap
40	10	1.5	194	23	US-09-950-083-3079	Sequence 3079, Ap
41	10	1.5	263	1	PCT-US01-27760-504	Sequence 504, App
42	10	1.5	263	1	PCT-US01-27760A-504	Sequence 504, App
43	10	1.5	263	21	US-09-790-893-3	Sequence 6, App1
44	10	1.5	284	16	US-09-056-019-3	Sequence 3, App1
45	10	1.5	284	16	US-09-287-788-3	Sequence 3, App1
46	10	1.5	406	14	US-09-056-019-1	Sequence 1, App1
47	10	1.5	406	14	US-09-287-788-1	Sequence 1, App1
48	10	1.5	428	16	US-09-056-019-24	Sequence 24, App1
49	10	1.5	428	16	US-09-287-788-24	Sequence 24, App1
50	10	1.5	428	20	US-09-605-7038-1256	Sequence 1256, Ap
51	10	1.5	428	20	US-09-605-7038-1258	Sequence 1258, Ap
52	10	1.5	438	21	US-09-738-626-6703	Sequence 6703, Ap
53	10	1.5	446	16	US-09-286-9818-6	Sequence 6, App1
54	10	1.5	446	16	US-09-286-9818-9	Sequence 9, App1
55	10	1.5	459	14	US-09-056-019-39	Sequence 39, App1
56	10	1.5	459	16	US-09-287-788-39	Sequence 39, App1
57	10	1.5	460	14	US-09-056-019-38	Sequence 38, App1
58	10	1.5	460	16	US-09-287-788-38	Sequence 38, App1
59	10	1.5	462	24	US-10-091-007-92	Sequence 92, App1
60	10	1.5	511	16	US-09-298-523B-13	Sequence 13, App1
61	10	1.5	511	16	US-09-298-523C-13	Sequence 13, App1
62	10	1.5	511	21	US-09-748-875-13	Sequence 13, App1
63	10	1.5	513	16	US-09-298-523B-12	Sequence 12, App1
64	10	1.5	513	16	US-09-298-523C-12	Sequence 12, App1
65	10	1.5	513	21	US-09-748-875-12	Sequence 12, App1
66	10	1.5	655	14	US-09-056-019-2	Sequence 2, App1
67	10	1.5	655	16	US-09-287-788-2	Sequence 2, App1
68	10	1.5	693	21	US-09-769-787-185	Sequence 185, App
69	10	1.5	698	22	US-09-739-449-11825	Sequence 11825, A
70	10	1.5	698	22	US-09-803-110-11825	Sequence 11825, A
71	10	1.5	757	21	US-09-739-449-11939	Sequence 11939, A
72	10	1.5	757	22	US-09-803-110-11939	Sequence 11939, A
73	10	1.5	774	1	PCT-US02-03987-12046	Sequence 12046, A
74	10	1.5	774	19	US-09-543-681A-5851	Sequence 5851, Ap
75	10	1.5	774	22	US-09-815-242-12046	Sequence 12046, A
76	10	1.5	774	22	US-10-072-851-12046	Sequence 12046, A
77	10	1.5	777	16	US-09-252-991A-23697	Sequence 23697, A
78	10	1.5	781	1	PCT-US02-03987-11277	Sequence 11277, A
79	10	1.5	781	22	US-09-815-242-11277	Sequence 11277, A
80	10	1.5	781	24	US-10-072-851-11277	Sequence 11277, A

81	10	1.5	814	17	US-09-328-352-4373	Sequence 4373, Ap
82	10	1.5	825	22	US-09-897-516-7108	Sequence 7108, Ap
83	10	1.5	825	22	US-60-215-161-7108	Sequence 7108, Ap
84	10	1.5	1704	4	US-08-089-372-40	Sequence 40, Appl
85	10	1.5	1704	8	US-08-440-522-40	Sequence 40, Appl
86	10	1.5	1704	8	US-08-440-522a-40	Sequence 40, Appl
87	10	1.5	1704	8	US-08-485-355a-40	Sequence 40, Appl
88	10	1.5	1704	16	US-09-234-238-40	Sequence 40, Appl

ALIGNMENTS

RESULT

1

US-09-536-784-2

; Sequence 2, Application US/09536784

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/536,784

; FILING DATE: 30-Oct-1997

; CLASSIFICATION: <Unknown>

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/961,083

; FILING DATE: OCT-30-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Michelle S. Marks

; REGISTRATION NUMBER: 41,971

; REFERENCE/DOCKET NUMBER: PB340P3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 666 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-536-784-2

Query Match 100.0%; Score 666; DB 19; Length 666;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNOLIADGSRERVAQAANDIPTDLVKAIVSIEDHRRFDRHGIDTIRILGAFLRN 60
DB 1 KIYDNKNOLIADGSRERVAQAANDIPTDLVKAIVSIEDHRRFDRHGIDTIRILGAFLRN 60
QY 61 LOSNSLGGSTLTQOLIKLTYFSTSDQTSKRKAQEWLAIOLEKATKQEIILTYINK 120
DB 61 LOSNSLGGSTLTQOLIKLTYFSTSDQTSKRKAQEWLAIOLEKATKQEIILTYINK 120
QY 121 VYMSNGNYGMQTAANYGKDLNNLSLPOLALLAGMPQAPNOVDYPSHPAADDRRLVYL 180
DB 121 VYMSNGNYGMQTAANYGKDLNNLSLPOLALLAGMPQAPNOVDYPSHPAADDRRLVYL 180

QY 181 SEMKNQYISAEQYERKAVNTPTDGLQSLKSASNPAYMDNLYKEVINQVEETGYNLLT 240
DB 181 SEMKNQYISAEQYERKAVNTPTDGLQSLKSASNPAYMDNLYKEVINQVEETGYNLLT 240
QY 241 TGMVYTNVQOAKHLMIDYNTDEYAYPPDELOVASTIVDSNGKVIAOLGARHOSN 300
DB 241 TGMVYTNVQOAKHLMIDYNTDEYAYPPDELOVASTIVDSNGKVIAOLGARHOSN 300
QY 301 VSFGINQAVETNRDQSTMKPTIDYAPALEGYVDSTATIVHDEPYNYPGTNTPYVMWR 360
DB 301 VSFGINQAVETNRDQSTMKPTIDYAPALEGYVDSTATIVHDEPYNYPGTNTPYVMWR 360
QY 361 GYFGNITLQYALQOSRNVPAVETLNKGLNRARFTFLNGLIDYPSIHYSAISSNTTESD 420
DB 361 GYFGNITLQYALQOSRNVPAVETLNKGLNRARFTFLNGLIDYPSIHYSAISSNTTESD 420
QY 421 KKYGASSEKMAAAYAFANGTYKKPYIHKYVPSDSSEKESFNVCGRRAKETAYMTD 480
DB 421 KKYGASSEKMAAAYAFANGTYKKPYIHKYVPSDSSEKESFNVCGRRAKETAYMTD 480
QY 481 MKTIVLTGTRNAYLAMLPOAGKTGTSNTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 540
DB 481 MKTIVLTGTRNAYLAMLPOAGKTGTSNTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 540
QY 541 AVWTGYSNRLPLVGNGLVAAKAVYSMTYLSGNSPEDMNIPEGLYRNGEEVFKNGAR 600
DB 541 AVWTGYSNRLPLVGNGLVAAKAVYSMTYLSGNSPEDMNIPEGLYRNGEEVFKNGAR 600
QY 601 STWNSPAPQOPSTESSSSSDSSTSSSTSTNNSTTTNNNTQCSNTTTPDOONON 660
DB 601 STWNSPAPQOPSTESSSSSDSSTSSSTSTNNSTTTNNNTQCSNTTTPDOONON 660
QY 661 POPAP 666
DB 661 POPAP 666

RESULT

2

US-09-765-271-2

; Sequence 2, Application US/09765271

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/765,271

; FILING DATE: 22-Jan-2001

; CLASSIFICATION: <Unknown>

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 09/536,784

; FILING DATE: <Unknown>

; APPLICATION NUMBER: 08/961,083

; FILING DATE: OCT-30-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Michelle S. Marks

; REGISTRATION NUMBER: 41,971

; REFERENCE/DOCKET NUMBER: PB340P3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-765-271-2

Query Match 100.0%; Score 666; DB 21; Length 666;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTYDNKNOLIADLGSERRNAOANDIPTDLVKAIVSIEDHREFDHGIDITIRLGAFLN 60
DB 1 KTYDNKNOLIADLGSERRNAOANDIPTDLVKAIVSIEDHREFDHGIDITIRLGAFLN 60
OY 61 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEAMLAIQLEBKATKOEILTYIYNK 120
DB 61 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEAMLAIQLEBKATKOEILTYIYNK 120
OY 121 VYMSNGNTGMOTAAONYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHEEAQDRRLNVL 180
DB 121 VYMSNGNTGMOTAAONYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHEEAQDRRLNVL 180
OY 181 SEMKNOGYISAQOYERKAVNPTITDGLQSLKASNPAYMDNLYKEVINOVEETGYNLT 240
DB 181 SEMKNOGYISAQOYERKAVNPTITDGLQSLKASNPAYMDNLYKEVINOVEETGYNLT 240
OY 241 TGMDEVYTNDOAQKHLMIDYNTDEVVAYPDELOVASTIVDSNGKVIQAOLGARHOSN 300
DB 241 TGMDEVYTNDOAQKHLMIDYNTDEVVAYPDELOVASTIVDSNGKVIQAOLGARHOSN 300
OY 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGYVDSTATIVHDEPNYNGTNPVYNMNR 360
DB 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGYVDSTATIVHDEPNYNGTNPVYNMNR 360
OY 361 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFNLGIDYPSIHNSNAISSNTTESD 420
DB 361 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFNLGIDYPSIHNSNAISSNTTESD 420
OY 421 KKYGASSEKMAAAYAFANGTYYPMYIHKVFSDGSEKESNVGTRAMKETTYAYMMD 480
DB 421 KKYGASSEKMAAAYAFANGTYYPMYIHKVFSDGSEKESNVGTRAMKETTYAYMMD 480
OY 481 MKKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 540
DB 481 MKKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 540
OY 541 AVMTGYSNRLPVLVNGSLTVAAKYVRSMWTVLSEGSNPEDWMIPEGLVRNGSEFVRKNGR 600
DB 541 AVMTGYSNRLPVLVNGSLTVAAKYVRSMWTVLSEGSNPEDWMIPEGLVRNGSEFVRKNGR 600
OY 601 STWNSPAPQOPSTESSSSSDSTSSQSSSTTPSTNNSTTTPNNNTQOOSNTTPOQONN 660
DB 601 STWNSPAPQOPSTESSSSSDSTSSQSSSTTPSTNNSTTTPNNNTQOOSNTTPOQONN 660
OY 661 POPAOP 666
DB 661 POPAOP 666

RESULT 3
US-09-765-272-2
Sequence 2, Application US/09765272
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESSES:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue

CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-765-272-2

Query Match 100.0%; Score 666; DB 21; Length 666;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTYDNKNOLIADLGSERRNAOANDIPTDLVKAIVSIEDHREFDHGIDITIRLGAFLN 60
DB 1 KTYDNKNOLIADLGSERRNAOANDIPTDLVKAIVSIEDHREFDHGIDITIRLGAFLN 60
OY 61 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEAMLAIQLEBKATKOEILTYIYNK 120
DB 61 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEAMLAIQLEBKATKOEILTYIYNK 120
OY 121 VYMSNGNTGMOTAAONYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHEEAQDRRLNVL 180
DB 121 VYMSNGNTGMOTAAONYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHEEAQDRRLNVL 180
OY 181 SEMKNOGYISAQOYERKAVNPTITDGLQSLKASNPAYMDNLYKEVINOVEETGYNLT 240
DB 181 SEMKNOGYISAQOYERKAVNPTITDGLQSLKASNPAYMDNLYKEVINOVEETGYNLT 240
OY 241 TGMDEVYTNDOAQKHLMIDYNTDEVVAYPDELOVASTIVDSNGKVIQAOLGARHOSN 300
DB 241 TGMDEVYTNDOAQKHLMIDYNTDEVVAYPDELOVASTIVDSNGKVIQAOLGARHOSN 300
OY 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGYVDSTATIVHDEPNYNGTNPVYNMNR 360
DB 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGYVDSTATIVHDEPNYNGTNPVYNMNR 360
OY 361 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFNLGIDYPSIHNSNAISSNTTESD 420
DB 361 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFNLGIDYPSIHNSNAISSNTTESD 420
OY 421 KKYGASSEKMAAAYAFANGTYYPMYIHKVFSDGSEKESNVGTRAMKETTYAYMMD 480
DB 421 KKYGASSEKMAAAYAFANGTYYPMYIHKVFSDGSEKESNVGTRAMKETTYAYMMD 480
OY 481 MKKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 540
DB 481 MKKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 540


```

QY      541  AMVTCTSNRRLFTLVANGGLTAAKYYRSMATYLTSEGSNPDMMNIPGLTYNGEYFVFNKAR  600
      |||||
Db      541  AMVTCTSNRRLFTLVANGGLTAAKYYRSMATYLTSEGSNPDMMNIPGLTYNGEYFVFNKAR  600
      |||||

QY      601  STWNSPAPQPPSTESSSSSDSSTSOSSSTTPSTNNSTTTNPNNNTTQOOSNTTPOOQON  660
      |||||
Db      601  STWNSPAPQPPSTESSSSSDSSTSOSSSTTPSTNNSTTTNPNNNTTQOOSNTTPOOQON  660
      |||||

QY      661  POPAP  666
      |||||
Db      661  POPAP  666

RESULT      4
US-09-583-110-3973
: Sequence 3973, Application US/09583110
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al.
: TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
: FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
: CURRENT APPLICATION NUMBER: US/09/583,110
: CURRENT FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/107,433
: PRIOR FILING DATE: 1998-06-30
: PRIOR APPLICATION NUMBER: US 60/085,131
: PRIOR FILING DATE: 1998-05-12
: PRIOR APPLICATION NUMBER: US 60/051,553
: PRIOR FILING DATE: 1997-07-02
: NUMBER OF SEQ ID NOS: 5322
: SEQ ID NO 3973
: LENGTH: 719
: TYPE: PROT
: ORGANISM: Streptococcus pneumoniae
: US-09-583-110-3973

```

Query Match	84.8%	Score 565	DB 19	Length 719
Best Local Similarity	99.8%	Pred. No. 0		
Matches 663	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY 1 KIYDKNKLIDLCSERRVNAQANDIPFDLYVAIYSIEDHREFDRHGIDITRILGAFLEN 60				
Db 54 KIYDKNKLIDLCSERRVNAQANDIPFDLYVAIYSIEDHREFDRHGIDITRILGAFLEN 113				
QY 61 LOSNSLGGSTLTPOOLIKLTYFSTSTSDQTSIRKQEAAMLAIOLEOKATQETLTYIYNK 120				
Db 114 LOSNSLGGSTLTPOOLIKLTYFSTSTSDQTSIRKQEAAMLAIOLEOKATQETLTYIYNK 173				
QY 121 VYMSNGNNGMOTAAQANYGKDLNNLSLQALLAAGMPQAPNOYDPSYSHPEAADRNLVL 180				
Db 174 VYMSNGNNGMOTAAQANYGKDLNNLSLQALLAAGMPQAPNOYDPSYSHPEAADRNLVL 233				
QY 181 SEMKNQGYISAEQYKAVNTPITDGLQSLKASNPAYMDNLTKEYINQVEETGYNLLT 240				
Db 234 SEMKNQGYISAEQYKAVNTPITDGLQSLKASNPAYMDNLTKEYINQVEETGYNLLT 293				
QY 241 TGMDEVYTNVDOEAOKHLMDIYNTDEYVAYPDDLOVASTIYDVNSGKYAOLGARHOSN 300				
Db 294 TGMDEVYTNVDOEAOKHLMDIYNTDEYVAYPDDLOVASTIYDVNSGKYAOLGARHOSN 353				
QY 301 VSFQGNQAVENFRDMSGTMRPTTDYAPALAEVGYVYSTATIYHDEPYNPGTNTPYVNMDR 360				
Db 354 VSFQGNQAVENFRDMSGTMRPTTDYAPALAEVGYVYSTATIYHDEPYNPGTNTPYVNMDR 413				
QY 361 GYFGNITLQVALQOSRNPAYETLKKVGLNRKKTFLNGIGIDYPSIHSNALSSTSTESD 420				
Db 414 GYFGNITLQVALQOSRNPAYETLKKVGLNRKKTFLNGIGIDYPSIHSNALSSTSTESD 473				
QY 421 KKYGASSEKMAAAYAFANGGTYKPYMILHKVVFDSGSEKESNNGTMRMKTTPYVMMMD 480				
Db 474 KKYGASSEKMAAAYAFANGGTYKPYMILHKVVFDSGSEKESNNGTMRMKTTPYVMMMD 533				

QY	481	MMVLVLYGGGRNAYLAMLEPOAKGTGTSNTDEIENH1KTSOFAPDELFGYTRKTSM	540
Db	534	MMKVLVLYGGGRNAYLAMLEPOAKGTGTSTNTDEIENH1KTSOFAPDELFGYTRKTSM	593
QY	541	AVMTGYSNRLPLPLVNGNLTYAAKVFYSMTY1SEGSNPEDMNIPEGLYRNGEFVEKNGAR	600
Db	594	AVMTGYSNRLPLPLVNGNLTYAAKVFYSMTY1SEGSNPEDMNIPEGLYRNGEFVEKNGAR	653
QY	601	STWNSPAPQOPPTSESSSSSDSSTQSSSTTPSTNNSTTTPNNNTTQOOSNTTPOOQON	660
Db	654	STWNSPAPQOPPTSESSSSSDSSTQSSSTTPSTNNSTTTPNNNTTQOOSNTTPOOQON	713
QY	661	POPAOP 666	
Db	714	POPAOP 719	

```

1      RESULT      5
2      US-09-107-433-3705
3      : Sequence 3705, Application US/09107433
4      :
5      : GENERAL INFORMATION:
6      : APPLICANT: Lynn A Doucette-Stamm and David Bush
7      : TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
8      : THERAPEUTICS RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIA
9      : THERAPEUTICS
10     :
11     : NUMBER OF SEQUENCES: 5206
12     :
13     : CORRESPONDENCE ADDRESS:
14     : ADDRESSER: GENOME THERAPEUTICS CORPORATION
15     : STREET: 100 Beaver Street
16     : CITY: Waltham
17     : STATE: Massachusetts
18     : COUNTRY: USA
19     : ZIP: 02354
20     :
21     : COMPUTER READABLE FORM:
22     : MEDIUM TYPE: CD-ROM ISO9660
23     : COMPUTER: <Unknown>
24     : OPERATING SYSTEM: <Unknown>
25     : SOFTWARE: <Unknown>
26     :
27     : CURRENT APPLICATION DATA:
28     : APPLICATION NUMBER: US/09/107,433
29     : FILING DATE: 30-Jun-1998
30     :
31     : PRIOR APPLICATION DATA:
32     : APPLICATION NUMBER: 60/085131
33     : FILING DATE: May 12, 1998
34     : APPLICATION NUMBER: 60/051553
35     : FILING DATE: July 2, 1997
36     :
37     : ATTORNEY/AGENT INFORMATION:
38     : NAME: Ariadello, Pamela Deneke
39     : REGISTRATION NUMBER: 40,489
40     : REFERENCE/DOCKET NUMBER: GTC-011
41     : TELECOMMUNICATION INFORMATION:
42     : TELEPHONE: (781)893-5007
43     : TELEFAX: (781)893-8277
44     :
45     : INFORMATION FOR SEQ ID NO: 3705:
46     : SEQUENCE CHARACTERISTICS:
47     : LENGTH: 721 amino acids
48     : TYPE: amino acid
49     : TOPOLOGY: linear
50     : MOLECULE TYPE: protein
51     : HYPOTHEICAL: YES
52     :
53     : ORIGINAL SOURCE:
54     : ORGANISM: Streptococcus pneumoniae
55     :
56     : FEATURE:
57     : NAME/KEY: misc_feature
58     : LOCATION: (B) LOCATION 1...721
59     : SEQUENCE DESCRIPTION: SEQ ID NO: 3705:
60     :
61     : US-09-107-433-3705

```

Query Match	84.8%;	Score 565;	DB 15;	Length 721;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 665; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;


```

OY 1 KIYDNKNOIADLGSERRVNAQANDIPTDLVKAIVSIEDHRFPDHGIDTIRILGAFERN 60
DB 56 KIYDNKNOIADLGSERRVNAQANDIPTDLVKAIVSIEDHRFPDHGIDTIRILGAFERN 115
OY 61 LOSNSILOGSSTLTQOLIKLTYFSTSDOTISKRAOEAMLAIOLEOKATKOEILTYINK 120
DB 116 LOSNSILOGSSTLTQOLIKLTYFSTSDOTISKRAOEAMLAIOLEOKATKOEILTYINK 175
OY 121 VYMSNGYMOAONRYKCDLNNLSLPOLALLAGMPOAPNOVDPSHPRAADRNNLYL 180
DB 176 VYMSNGYMOAONRYKCDLNNLSLPOLALLAGMPOAPNOVDPSHPRAADRNNLYL 235
OY 181 SEKNNGYISAEQYKAVNTPITDGLQSLKASNPAYMDNYLKEVINOVEEFTGNLTL 240
DB 236 SEKNNGYISAEQYKAVNTPITDGLQSLKASNPAYMDNYLKEVINOVEEFTGNLTL 295
OY 241 TGMNDVYTNVDOEAKHLMIDYNTDEVYAYPDELOVASTIVDSNGKVIAOLGARHOSN 300
DB 296 TGMNDVYTNVDOEAKHLMIDYNTDEVYAYPDELOVASTIVDSNGKVIAOLGARHOSN 355
OY 301 VSEGINQAVETNRDMSYTKKPTIDYAPALEGYVDSTATIVHDEPNYNGTNPYNNMDR 360
DB 356 VSEGINQAVETNRDMSYTKKPTIDYAPALEGYVDSTATIVHDEPNYNGTNPYNNMDR 415
OY 361 GYFGNTTLOVALQOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHYNSAISSNTTESD 420
DB 416 GYFGNTTLOVALQOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHYNSAISSNTTESD 475
OY 421 KYGASSEKMAAAYAFANGGTYKPMYTHKVVFSDSGEKESNVGTRAMKETTAYMMTD 480
DB 476 KYGASSEKMAAAYAFANGGTYKPMYTHKVVFSDSGEKESNVGTRAMKETTAYMMTD 535
OY 481 MKKTYVTYGTGRNAYLAMPQAGKTGTSNVTDEIEHNHKTQSOFVAPDELFAGYTRKYSN 540
DB 536 MKKTYVTYGTGRNAYLAMPQAGKTGTSNVTDEIEHNHKTQSOFVAPDELFAGYTRKYSN 595
OY 541 AVWTGYSNRLTPLYNGLTVAAKYVRSMTYLSGSPNEDWNIPEGLYRNGEYFKNRGAR 600
DB 596 AVWTGYSNRLTPLYNGLTVAAKYVRSMTYLSGSPNEDWNIPEGLYRNGEYFKNRGAR 655
OY 601 STWNSPAPQOPSTESSSSSSSSSTSSQSSSTPTNNSTTTNNNTTQOOSNTTPOOONON 660
DB 656 STWNSPAPQOPSTESSSSSSSSSTSSQSSSTPTNNSTTTNNNTTQOOSNTTPOOONON 715
OY 661 POPAP 666
DB 716 POPAP 721

```

RESULT 6

```

PCT-US02-03987-13423
; Sequence 13423, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits
; FILE REFERENCE: ELITRA 028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
PCT-US02-03987-13423

```

Query Match 49.5%; Score 330; DB 1; Length 719;
Best Local Similarity 99.6%; Pred. No. 0;

Matches 530; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 72 LFOOLIKLTYFSTSDOTSRRKAOEAMLAIOLEOKATKOEILTYINKYMSNGNGMO 131
DB 125 LFOOLIKLTYFSTSDOTSRRKAOEAMLAIOLEOKATKOEILTYINKYMSNGNGMO 184
OY 132 TAAQNYGKDLNNLSLPOLALLAGMPOAPNOVDPSHPRAADRNNLYLSEKNNGYISA 191
DB 185 TAAQNYGKDLNNLSLPOLALLAGMPOAPNOVDPSHPRAADRNNLYLSEKNNGYISA 244
OY 192 EYKAVNTPITDGLQSLKASNPAYMDNYLKEVINOVEEFTGNLTLTGMNDVYTNVDO 251
DB 245 EYKAVNTPITDGLQSLKASNPAYMDNYLKEVINOVEEFTGNLTLTGMNDVYTNVDO 304
OY 252 EAKHLMIDYNTDEVYAYPDELOVASTIVDSNGKVIAOLGARHOSNVSFGINQAVET 311
DB 305 EAKHLMIDYNTDEVYAYPDELOVASTIVDSNGKVIAOLGARHOSNVSFGINQAVET 364
OY 312 NRDMGSTMKPTIDYAPALEGYVDSTATIVHDEPNYNGTNPYNNMDRGYFGNTTLOVA 371
DB 365 NRDMGSTMKPTIDYAPALEGYVDSTATIVHDEPNYNGTNPYNNMDRGYFGNTTLOVA 424
OY 372 LOOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHYNSAISSNTTESDCKKYGASSEKMA 431
DB 425 LOOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHYNSAISSNTTESDCKKYGASSEKMA 484
OY 432 AAYAAFANGGTYKPMYTHKVVFSDSGEKESNVGTRAMKETTAYMMTDMMKTYVTYGTG 491
DB 485 AAYAAFANGGTYKPMYTHKVVFSDSGEKESNVGTRAMKETTAYMMTDMMKTYVTYGTG 544
OY 492 RNAVYLA MLPQAGKTGTSNVTDEIEHNHKTQSOFVAPDELFAGYTRKYSNRLT 551
DB 545 RNAVYLA MLPQAGKTGTSNVTDEIEHNHKTQSOFVAPDELFAGYTRKYSNRLT 604
OY 552 PLYNGLTVAAKYVRSMTYLSGSPNEDWNIPEGLYRNGEYFKNRGARSTW 603
DB 605 PLYNGLTVAAKYVRSMTYLSGSPNEDWNIPEGLYRNGEYFKNRGARSTW 656

```

RESULT 7

```

US-09-815-242-13423
; Sequence 13423, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Olsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: Prokaryotes
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423

```


LENGTH: 719
TYPE: PR
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13423

Query Match 49.5%; Score 330; DB 22; Length 719;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 530; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 72 LTQOLIKLTFSTSTQSTSRKAQEAFLAIQLEOKATKOEILLYINKYMSNGNGM 131
DB 125 LTQOLIKLTFSTSTQSTSRKAQEAFLAIQLEOKATKOEILLYINKYMSNGNGM 184
QY 132 TAAQNTYGGKDLNLSLPOLALLAGMPAPNOYDPSHPPEAODRRNLVLSMKNOGYISA 191
DB 185 TAAQNTYGGKDLNLSLPOLALLAGMPAPNOYDPSHPPEAODRRNLVLSMKNOGYISA 244
QY 192 EGYEKAQNTPTTDLQSLKASNPAYMDNYLKEVINQVEETGYNLLTGMGYTVNDQ 251
DB 245 EGYEKAQNTPTTDLQSLKASNPAYMDNYLKEVINQVEETGYNLLTGMGYTVNDQ 304
QY 252 EAQKHLMDIYNTDEYVAPDELOVASTIVDSNGKYIAQIGARHOSNVSFGINQAVET 311
DB 305 EAQKHLMDIYNTDEYVAPDELOVASTIVDSNGKYIAQIGARHOSNVSFGINQAVET 364
QY 312 NRDMGSTMKPITDYAPALEGYVDSTATIVHDEPYNPGTNPVYNNDRGYFGNITLOYA 371
DB 365 NRDMGSTMKPITDYAPALEGYVDSTATIVHDEPYNPGTNPVYNNDRGYFGNITLOYA 424
QY 372 LOOSRNPAYETLNKVLNRAKTFPLNGLGIDYPSIHYSNAISSNTTESDKKYGASSEMA 431
DB 425 LOOSRNPAYETLNKVLNRAKTFPLNGLGIDYPSIHYSNAISSNTTESDKKYGASSEMA 484
QY 432 AAYAAAFANGSTGYKPMYIHKVVFSDGSEKEFSNNGTRAMETAYMMTDMMKTVLYGTG 491
DB 485 AAYAAAFANGSTGYKPMYIHKVVFSDGSEKEFSNNGTRAMETAYMMTDMMKTVLYGTG 544
QY 492 RNAYLAWLPQAGKTGTSNTYDEIEENHIKTSQFAPADELAGYTRKXSMVWVGYSNRLT 551
DB 545 RNAYLAWLPQAGKTGTSNTYDEIEENHIKTSQFAPADELAGYTRKXSMVWVGYSNRLT 604
QY 552 PLYVNGILTVAAKYRSMATYLSGSPEDWNIPBGLYRNGEFYFKNGARSTW 603
DB 605 PLYVNGILTVAAKYRSMATYLSGSPEDWNIPBGLYRNGEFYFKNGARSTW 656

RESULT 8
US-10-072-851-13423

Sequence 13423, Application US/10072851
GENERAL INFORMATION:

APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits
FILE REFERENCE: ELITRA.028A
CURRENT APPLICATION NUMBER: US/10/072,851
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13423
LENGTH: 719
TYPE: PR
ORGANISM: Streptococcus pneumoniae
US-10-072-851-13423

Query Match 49.5%; Score 330; DB 24; Length 719;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 530; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 72 LTQOLIKLTFSTSTQSTSRKAQEAFLAIQLEOKATKOEILLYINKYMSNGNGM 131
DB 125 LTQOLIKLTFSTSTQSTSRKAQEAFLAIQLEOKATKOEILLYINKYMSNGNGM 184
QY 132 TAAQNTYGGKDLNLSLPOLALLAGMPAPNOYDPSHPPEAODRRNLVLSMKNOGYISA 191
DB 185 TAAQNTYGGKDLNLSLPOLALLAGMPAPNOYDPSHPPEAODRRNLVLSMKNOGYISA 244
QY 192 EGYEKAQNTPTTDLQSLKASNPAYMDNYLKEVINQVEETGYNLLTGMGYTVNDQ 251
DB 245 EGYEKAQNTPTTDLQSLKASNPAYMDNYLKEVINQVEETGYNLLTGMGYTVNDQ 304
QY 252 EAQKHLMDIYNTDEYVAPDELOVASTIVDSNGKYIAQIGARHOSNVSFGINQAVET 311
DB 305 EAQKHLMDIYNTDEYVAPDELOVASTIVDSNGKYIAQIGARHOSNVSFGINQAVET 364
QY 312 NRDMGSTMKPITDYAPALEGYVDSTATIVHDEPYNPGTNPVYNNDRGYFGNITLOYA 371
DB 365 NRDMGSTMKPITDYAPALEGYVDSTATIVHDEPYNPGTNPVYNNDRGYFGNITLOYA 424
QY 372 LOOSRNPAYETLNKVLNRAKTFPLNGLGIDYPSIHYSNAISSNTTESDKKYGASSEMA 431
DB 425 LOOSRNPAYETLNKVLNRAKTFPLNGLGIDYPSIHYSNAISSNTTESDKKYGASSEMA 484
QY 432 AAYAAAFANGSTGYKPMYIHKVVFSDGSEKEFSNNGTRAMETAYMMTDMMKTVLYGTG 491
DB 485 AAYAAAFANGSTGYKPMYIHKVVFSDGSEKEFSNNGTRAMETAYMMTDMMKTVLYGTG 544
QY 492 RNAYLAWLPQAGKTGTSNTYDEIEENHIKTSQFAPADELAGYTRKXSMVWVGYSNRLT 551
DB 545 RNAYLAWLPQAGKTGTSNTYDEIEENHIKTSQFAPADELAGYTRKXSMVWVGYSNRLT 604
QY 552 PLYVNGILTVAAKYRSMATYLSGSPEDWNIPBGLYRNGEFYFKNGARSTW 603
DB 605 PLYVNGILTVAAKYRSMATYLSGSPEDWNIPBGLYRNGEFYFKNGARSTW 656

RESULT 9
US-08-116-541-4

Sequence 4, Application US/08116541
GENERAL INFORMATION:

APPLICANT: Masure, H. Robert
APPLICANT: Pearce, Barbara J.
APPLICANT: Toumanen, Elaine
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/116,541

FILING DATE: 19930901
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-069
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-116-541-4

Query Match 36.8%; Score 245; DB 5; Length 320;
Best Local Similarity 100.0%; Pred. No. 1.6e-232; Indels 0; Gaps 0;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 LTQQLIKLTFSTSTSDQTSRKAQEWLAIQLEKATKQELITYINKVYNSNGNYGMQ 131
DB 76 LTQQLIKLTFSTSTSDQTSRKAQEWLAIQLEKATKQELITYINKVYNSNGNYGMQ 135

QY 132 TAAQNTYGGDNLNLSLPQALLAGMPQAPNOYDPYSHPEAQRRLVYSEKNOGYISA 191
DB 136 TAAQNTYGGDNLNLSLPQALLAGMPQAPNOYDPYSHPEAQRRLVYSEKNOGYISA 195

QY 192 EGYEKAVNPTITDGLQSLKASNYPAYMDNYLKEVINQVEETGYMLTTGMDVYTNVDO 251
DB 196 EGYEKAVNPTITDGLQSLKASNYPAYMDNYLKEVINQVEETGYMLTTGMDVYTNVDO 255

QY 252 EAQKHLMDIYNTDEVYAYPDELQVASTIVDSNGKVIQAQLGARHSSNVSEGINQAVET 311
DB 256 EAQKHLMDIYNTDEVYAYPDELQVASTIVDSNGKVIQAQLGARHSSNVSEGINQAVET 315

QY 312 NRDMG 316
DB 316 NRDMG 320

RESULT 10
US-08-116-541-24
Sequence 24, Application US/08116541
GENERAL INFORMATION:
APPLICANT: Masure, H. Robert
APPLICANT: Pearce, Barbara J.
APPLICANT: Toumanou, Elaine
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
TITLE OF INVENTION: ACCELLULAR VACCINES BASED THERON
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/116,541
FILING DATE: 19930901
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-069
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
STRAIN: R6
IMMEDIATE SOURCE:
CLONE: SPRU42
US-08-116-541-24

Query Match 11.6%; Score 77; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3.6e-67;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 TTGMDVYTNVDOEAQKHLMDIYNTDEVYAYPDELQVASTIVDSNGKVIQAQLGARHSS 299
DB 1 TTGMDVYTNVDOEAQKHLMDIYNTDEVYAYPDELQVASTIVDSNGKVIQAQLGARHSS 60

QY 300 NVSEGINQAVETNRDMG 316
DB 61 NVSEGINQAVETNRDMG 77

RESULT 11
US-09-107-532-5667
Sequence 5667, Application US/09107532
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7308
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER:
OPERATING SYSTEM:
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532
FILING DATE:
APPLICATION DATA:
APPLICATION NUMBER: 60/085598
FILING DATE: May 14, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Denke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5667:

SEQUENCE CHARACTERISTICS:
LENGTH: 823 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1...823
US-09-107-532-5667

Query Match 2.7%; Score 18; DB 15; Length 823;
Best Local Similarity 100.0%; Pred. No. 6.6e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 429 KMAAAYAFANGTYRKP 446
Db 543 KMAAAYAFANGTYRKP 560

RESULT 12
US-09-107-532A-5667
Sequence 5667, Application US/09107532A
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinifello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-8277
TELEFAX: (781)893-5007
INFORMATION FOR SEQ ID NO: 5667:
SEQUENCE CHARACTERISTICS:
LENGTH: 823 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc.feature
LOCATION: (b) LOCATION 1...823
SEQUENCE DESCRIPTION: SEQ ID NO: 5667:
US-09-107-532A-5667

Query Match 2.7%; Score 18; DB 15; Length 823;
Best Local Similarity 100.0%; Pred. No. 6.6e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 429 KMAAAYAFANGTYRKP 446
Db 543 KMAAAYAFANGTYRKP 560

RESULT 13
US-09-634-238-229
Sequence 229, Application US/09634238
GENERAL INFORMATION:
APPLICANT: Glenn, Matthew
APPLICANT: Havukkala, Ilkka J.
APPLICANT: Bloksberg, Leonard, N.
APPLICANT: Lubbers, Mark W.
APPLICANT: Dekker, James
APPLICANT: Christensson, Anna C.
APPLICANT: Holland, Ross
APPLICANT: O'Toole, Paul W.
APPLICANT: Reid, Julian R.
APPLICANT: Coolbear, Timothy
TITLE OF INVENTION: Polynucleotides, materials incorporating
TITLE OF INVENTION: them and methods for using them.
FILE REFERENCE: 11000.104301
CURRENT APPLICATION NUMBER: US/09/634,238
CURRENT FILING DATE: 2000-08-08
NUMBER OF SEQ ID NOS: 422
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 229
LENGTH: 771
TYPE: PRT
ORGANISM: Lactobacillus rhamnosus
US-09-634-238-229

Query Match 2.1%; Score 14; DB 20; Length 771;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 LOGGSTLTQOLIKL 79
Db 143 LOGGSTLTQOLIKL 156

RESULT 14
US-09-634-238-262
Sequence 262, Application US/09634238
GENERAL INFORMATION:
APPLICANT: Glenn, Matthew
APPLICANT: Havukkala, Ilkka J.
APPLICANT: Bloksberg, Leonard, N.
APPLICANT: Lubbers, Mark W.
APPLICANT: Dekker, James
APPLICANT: Christensson, Anna C.
APPLICANT: Holland, Ross
APPLICANT: O'Toole, Paul W.
APPLICANT: Reid, Julian R.
APPLICANT: Coolbear, Timothy
TITLE OF INVENTION: Polynucleotides, materials incorporating
TITLE OF INVENTION: them and methods for using them.
FILE REFERENCE: 11000.104301
CURRENT APPLICATION NUMBER: US/09/634,238
CURRENT FILING DATE: 2000-08-08
NUMBER OF SEQ ID NOS: 422
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 262
LENGTH: 771
TYPE: PRT
ORGANISM: Lactobacillus rhamnosus
US-09-634-238-262


```
Query Match          2.1%; Score 14; DB 20; Length 771;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 LOGGSTLTQQLIKL 79
   |||||
DB 143 LOGGSTLTQQLIKL 156

RESULT 15
PCT-US02-03987-10728
; Sequence 10728, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; FILE REFERENCE: ELITTRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10728
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
PCT-US02-03987-10728

Query Match          2.1%; Score 14; DB 1; Length 778;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 LOGGSTLTQQLIKL 79
   |||||
DB 146 LOGGSTLTQQLIKL 159

RESULT 16
US-09-815-242-10728
; Sequence 10728, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITTRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10728
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10728

Query Match          2.1%; Score 14; DB 22; Length 778;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 LOGGSTLTQQLIKL 79
   |||||
DB 146 LOGGSTLTQQLIKL 159

RESULT 17
US-10-072-851-10728
; Sequence 10728, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit
; FILE REFERENCE: ELITTRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10728
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-10-072-851-10728

Query Match          2.1%; Score 14; DB 24; Length 778;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 LOGGSTLTQQLIKL 79
   |||||
DB 146 LOGGSTLTQQLIKL 159

RESULT 18
US-09-134-000-4939
; Sequence 4939, Application US/09134000A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS
; FILE REFERENCE: GTC-005
; CURRENT APPLICATION NUMBER: US/09/134,000A
; CURRENT FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 6810
; SEQ ID NO 4939
; LENGTH: 789
```


TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000-4939

Query Match 2.1%; Score 14; DB 15; Length 789;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 GGSTLTQOLIK 79
DB 157 GGSTLTQOLIK 170

RESULT 19
PCT-US97-14436-535
Sequence 535, Application PC/TUS9714436
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
NUMBER OF SEQUENCES: 713
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/14436
FILING DATE: 15-AUG-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,022
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50533
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 535:
SEQUENCE CHARACTERISTICS:
LENGTH: 266 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US97-14436-535

Query Match 1.7%; Score 11; DB 1; Length 266;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSTLTQOLIK 78
DB 107 GGSTLTQOLIK 117

RESULT 20
US-08-911-503-535
Sequence 535, Application US/08911503
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
NUMBER OF SEQUENCES: 713
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,503
FILING DATE: 15-AUG-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,022
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50533
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 535:
SEQUENCE CHARACTERISTICS:
LENGTH: 266 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-911-503-535

Query Match 1.7%; Score 11; DB 13; Length 266;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSTLTQOLIK 78
DB 107 GGSTLTQOLIK 117

RESULT 21
US-08-911-503A-535
Sequence 535, Application US/08911503A
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
NUMBER OF SEQUENCES: 713
CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,503A
FILING DATE: 15-AUG-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,022
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50533
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 535:
SEQUENCE CHARACTERISTICS:
LENGTH: 266 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-911-503A-535

Query Match 1.7%; Score 11; DB 13; Length 266;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 GGSITLQOLIK 78
|||||
DB 107 GGSITLQOLIK 117

RESULT 22
US-09-417-507-35782
Sequence 35782, Application US/09417507
GENERAL INFORMATION:
APPLICANT: KEITH G. WEINSTOCK ET AL.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
FILE REFERENCE: PAT99-10
CURRENT APPLICATION NUMBER: US/09/417,507
CURRENT FILING DATE: 1999-10-14
NUMBER OF SEQ ID NOS: 44312
SEQ ID NO 35782
LENGTH: 314
TYPE: PRT
ORGANISM: A.fumigatus
US-09-417-507-35782

Query Match 1.7%; Score 11; DB 18; Length 314;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSSTS 626
|||||
DB 148 SSSSSSDSSTS 158

RESULT 23

US-09-583-110-4219
Sequence 4219, Application US/09583110
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococci
FILE REFERENCE: PATH00-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 4219
LENGTH: 820
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-4219

Query Match 1.7%; Score 11; DB 19; Length 820;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 GGSITLQOLIK 78
|||||
DB 176 GGSITLQOLIK 186

RESULT 24
US-09-377-465-2
Sequence 2, Application US/09377465
GENERAL INFORMATION:
APPLICANT: Hoskins, JoAnn
APPLICANT: Jaskunas, S. Richard
APPLICANT: Zhao, Genshi
APPLICANT: Rockey, Pamela
TITLE OF INVENTION: Novel Penicillin Binding Protein From Streptococcus
FILE REFERENCE: X12498
CURRENT APPLICATION NUMBER: US/09/377,465
CURRENT FILING DATE: 1999-08-19
EARLIER APPLICATION NUMBER: 60/100,887
EARLIER FILING DATE: 1998-09-23
EARLIER APPLICATION NUMBER: 60/111,862
EARLIER FILING DATE: 1998-12-11
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 821
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-377-465-2

Query Match 1.7%; Score 11; DB 17; Length 821;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 GGSITLQOLIK 78
|||||
DB 176 GGSITLQOLIK 186

RESULT 25
US-09-377-465A-2
Sequence 2, Application US/09377465A
GENERAL INFORMATION:
APPLICANT: Hoskins, JoAnn
APPLICANT: Jaskunas, Stanley R

APPLICANT: Zhao, Genshi
APPLICANT: Rockey, Pamela K
TITLE OF INVENTION: NOVEL PENICILLIN BINDING PROTEIN FROM STREPTOCOCCUS
TITLE OF INVENTION: PNEUMONIAE
FILE REFERENCE: X12498 Sequence List
CURRENT APPLICATION NUMBER: US/09/377,465A
CURRENT FILING DATE: 1999-08-19
PRIOR APPLICATION NUMBER: 60/100,887
PRIOR FILING DATE: 1998-09-23
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 821
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-377-465A-2

Query Match 1.7%; Score 11; DB 17; Length 821;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 68 GGSSTLQOLIK 78
|||||
DB 176 GGSSTLQOLIK 186

RESULT 26
US-09-107-433-4760
Sequence 4760, Application US/09107433
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4760:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:

NAME/KEY: misc-feature
LOCATION: (B) LOCATION 1...834
SEQUENCE DESCRIPTION: SEQ ID NO: 4760:
US-09-107-433-4760

Query Match 1.7%; Score 11; DB 15; Length 834;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 68 GGSSTLQOLIK 78
|||||
DB 190 GGSSTLQOLIK 200

RESULT 27
US-09-621-976-7706
Sequence 7706, Application US/09621976
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 7706
LENGTH: 76
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 72
OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr
US-09-621-976-7706

Query Match 1.5%; Score 10; DB 20; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 616 SSSSSSDSST 625
|||||
DB 23 SSSSSSDSST 32

RESULT 28
US-09-834-366-24929
Sequence 24929, Application US/09834366
GENERAL INFORMATION:
APPLICANT: Bejani, Stephane
APPLICANT: Tanaka, Hiroaki
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Jobert, Severin
APPLICANT: Giordano, Jean-Yves
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: 81.US2.REG
CURRENT APPLICATION NUMBER: US/09/834,366
CURRENT FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/197,873
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 52153
SOFTWARE: Patent.pm
SEQ ID NO 24929
LENGTH: 76
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 72
OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr

US-09-834-366-24929

Query Match 1.5%; Score 10; DB 22; Length 76;

Best Local Similarity 100.0%; Pred. No. 0.55; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625

DB 23 SSSSSSDSST 32

RESULT 29

US-09-834-366-24930

; Sequence 24930, Application US/09834366

; GENERAL INFORMATION:

; APPLICANT: Bejain, Stephane

; APPLICANT: Tanaka, Hiroaki

; APPLICANT: Dumas Milne Edwards, Jean Baptiste

; APPLICANT: Jobert, Severin

; APPLICANT: Giordano, Jean-Yves

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: 81.052.REG

; CURRENT APPLICATION NUMBER: US/09/834.366

; CURRENT FILING DATE: 2001-04-13

; PRIOR APPLICATION NUMBER: US 60/197,873

; PRIOR FILING DATE: 2000-04-18

; NUMBER OF SEQ ID NOS: 52153

; SOFTWARE: Patent.pm

; SEQ ID NO 24930

; LENGTH: 76

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: 72

; OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr

US-09-834-366-24930

Query Match 1.5%; Score 10; DB 22; Length 76;

Best Local Similarity 100.0%; Pred. No. 0.55; Mismatches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625

DB 23 SSSSSSDSST 32

RESULT 30

US-60-147-499-7706

; Sequence 7706, Application US/60147499

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/60/147,499

; CURRENT FILING DATE: 1999-08-05

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 7706

; LENGTH: 76

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: 72

; OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr

US-60-147-499-7706

Query Match 1.5%; Score 10; DB 26; Length 76;

Best Local Similarity 100.0%; Pred. No. 0.55; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625

DB 23 SSSSSSDSST 32

RESULT 31

US-60-197-873-24929

; Sequence 24929, Application US/60197873

; GENERAL INFORMATION:

; APPLICANT: Bejain, Stephane

; APPLICANT: Tanaka, Hiroaki

; APPLICANT: Dumas Milne Edwards, Jean Baptiste

; APPLICANT: Jobert, Severin

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: 81.US1.PRO

; CURRENT APPLICATION NUMBER: US/60/197,873

; CURRENT FILING DATE: 2000-04-18

; NUMBER OF SEQ ID NOS: 52153

; SOFTWARE: Patent.pm

; SEQ ID NO 24929

; LENGTH: 76

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: 72

; OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr

US-60-197-873-24929

Query Match 1.5%; Score 10; DB 26; Length 76;

Best Local Similarity 100.0%; Pred. No. 0.55; Mismatches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625

DB 23 SSSSSSDSST 32

RESULT 32

US-60-197-873-24930

; Sequence 24930, Application US/60197873

; GENERAL INFORMATION:

; APPLICANT: Bejain, Stephane

; APPLICANT: Tanaka, Hiroaki

; APPLICANT: Dumas Milne Edwards, Jean Baptiste

; APPLICANT: Jobert, Severin

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: 81.US1.PRO

; CURRENT APPLICATION NUMBER: US/60/197,873

; CURRENT FILING DATE: 2000-04-18

; NUMBER OF SEQ ID NOS: 52153

; SOFTWARE: Patent.pm

; SEQ ID NO 24930

; LENGTH: 76

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: 72

; OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr

US-60-197-873-24930

Query Match 1.5%; Score 10; DB 26; Length 76;

Best Local Similarity 100.0%; Pred. No. 0.55; Mismatches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
Db 23 SSSSSSDSST 32

RESULT 33

US-09-536-784-102
; Sequence 102, Application US/09536784

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/536,784

FILING DATE: 30-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/961,083

FILING DATE: OCT-30-1997

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB340P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:

LENGTH: 110 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 102:

US-09-536-784-102

Query Match 1.5%; Score 10; DB 19; Length 110;

Best Local Similarity 100.0%; Pred. No. 0.78;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625

Db 18 SSSSSSDSST 27

RESULT 34

US-09-765-271-102
; Sequence 102, Application US/09765271

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/765,271

FILING DATE: 22-Jan-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/536,784

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/961,083

FILING DATE: OCT-30-1997

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB340P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:

LENGTH: 110 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 102:

US-09-765-271-102

Query Match 1.5%; Score 10; DB 21; Length 110;

Best Local Similarity 100.0%; Pred. No. 0.78;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625

Db 18 SSSSSSDSST 27

RESULT 35

US-09-765-272-102
; Sequence 102, Application US/09765272

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/765,272

FILING DATE: 22-Jan-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/961,083

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 102:
US-09-765-272-102

Query Match 1.5%; Score 10; DB 21; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
|||||
Db 18 SSSSSSDSST 27

RESULT 36
US-09-614-150-34311
Sequence 34311, Application US/09614150
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
FILE REFERENCE: CLO00728
CURRENT APPLICATION NUMBER: US/09/614,150
CURRENT FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637
PRIOR FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 43008
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 34311
LENGTH: 119
TYPE: PRT
ORGANISM: DROSOPHILA
US-09-614-150-34311

Query Match 1.5%; Score 10; DB 20; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.84;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 617 SSSSSSDSST 626
|||||
Db 60 SSSSSSDSST 69

RESULT 37
US-60-191-637-33895
Sequence 33895, Application US/60191637
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING

TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CLO00392
CURRENT APPLICATION NUMBER: US/60/191,637
CURRENT FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 42660
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 33895
LENGTH: 119
TYPE: PRT
ORGANISM: DROSOPHILA
US-60-191-637-33895

Query Match 1.5%; Score 10; DB 26; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.84;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 617 SSSSSSDSST 626
|||||
Db 60 SSSSSSDSST 69

RESULT 38
US-09-621-976-5866
Sequence 5866, Application US/09621976
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTS and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 5866
LENGTH: 130
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -36...-1
NAME/KEY: UNSURE
LOCATION: 90
OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr
US-09-621-976-5866

Query Match 1.5%; Score 10; DB 20; Length 130;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
|||||
Db 77 SSSSSSDSST 86

RESULT 39
US-60-147-499-5866
Sequence 5866, Application US/60147499
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTS and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/60/147,499
CURRENT FILING DATE: 1999-08-05
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 5866
LENGTH: 130

TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -36..-1
FEATURE:
NAME/KEY: UNSURE
LOCATION: 90
OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr
US-60-147-499-5866

Query Match 1.5%; Score 10; DB 26; Length 130;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
DB 77 SSSSSSDSST 86

RESULT 40
US-09-950-083-3079
Sequence 3079, Application US/09950083
GENERAL INFORMATION:
APPLICANT: Rosen, et. al
TITLE OF INVENTION: Human Secreted Proteins
FILE REFERENCE: PS805
CURRENT APPLICATION NUMBER: US/09/950, 083
CURRENT FILING DATE: 2001-09-12
PRIOR APPLICATION NUMBER: 60/278,650
PRIOR FILING DATE: 2001-03-27
PRIOR APPLICATION NUMBER: 09/833,245
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: PCT/US01/11988
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: PCT/US00/06043
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06012
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06058
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06044
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06059
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06042
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06014
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06013
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06049
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06057
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06824
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06765
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06792
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06830
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06782
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06822
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06791
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06828
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06823

PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06781
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/07505
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07440
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07506
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07507
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07535
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07525
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07534
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07483
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07526
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07527
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07661
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07579
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07723
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07724
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/14929
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/07722
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07578
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07726
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07677
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07725
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/09070
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08982
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08983
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09067
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09066
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09068
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08981
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08980
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09071
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09069
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/15136
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14926
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14963
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/15135
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14934
PRIOR FILING DATE: 2000-06-01

; PRIOR APPLICATION NUMBER: PCT/US00/14933
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/15137
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/14928
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/14973
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/14964
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/26376
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US00/26371
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US00/26324
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US00/26323
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US00/26337
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US01/13318
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/124,146
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/167,061
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/124,093
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/166,989
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/124,145
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/168,654
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,099
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/168,661
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,096
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/168,622
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,143
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/168,663
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,095
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/138,598
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: US 60/168,665
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/125,360
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: US 60/138,626
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: US 60/168,662
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,144
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/138,574
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: US 60/168,667
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,142
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/138,597
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: US 60/168,666
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/125,359
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: US 60/168,664

; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/126,051
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/169,906

Query Match 1.5%; Score 10; DB 23; Length 194;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
Db 149 SSSSSSDSST 158

RESULT 41
PCT-US01-27760-504
; Sequence 504, Application PC/TUS0127760
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-115/785
; CURRENT APPLICATION NUMBER: PCT/US01/27760
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/687,527
; PRIOR FILING DATE: 2000-10-12
; SOFTWARE: Custom
; SEQ ID NO 504
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-27760-504

Query Match 1.5%; Score 10; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
Db 218 SSSSSSDSST 227

RESULT 42
PCT-US01-27760A-504
; Sequence 504, Application PC/TUS0127760A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-115/785
; CURRENT APPLICATION NUMBER: PCT/US01/27760A
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/687,527
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 892
; SOFTWARE: Custom
; SEQ ID NO 504
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-27760A-504

Query Match 1.5%; Score 10; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
Db 218 SSSSSSDSST 227

RESULT 43

US-09-790-893-6
; Sequence 6, Application US/09790893
; GENERAL INFORMATION:
; APPLICANT: Mauch, Kimberly
; APPLICANT: Bush, Angie
; APPLICANT: Heichman, Karen
; APPLICANT: Roch, Jean-Marc
; APPLICANT: Bartel, Paul L.
; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
; FILE REFERENCE: 2318-293
; CURRENT APPLICATION NUMBER: US/09/790,893
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/185,080
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-790-893-6

Query Match 1.5%; Score 10; DB 21; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
Db 218 SSSSSSDSST 227

RESULT 44
US-09-056-019-3
; Sequence 3, Application US/09056019A
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Wizemann, Theresa
; APPLICANT: Masure, H. R.
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Koenig, Scott
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; FILE REFERENCE: 1340-1-017
; CURRENT APPLICATION NUMBER: US/09/056,019A
; CURRENT FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-056-019-3

Query Match 1.5%; Score 10; DB 14; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
Db 113 SSSSSSDSST 122

RESULT 45
US-09-287-788-3
; Sequence 3, Application US/09287788B
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Masure, H. R.
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL

; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; TITLE OF INVENTION: THEREFROM AND USES THEREOF
; FILE REFERENCE: 1340-1-017N
; CURRENT APPLICATION NUMBER: US/09/287,788B
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: 60/080,878
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-287-788-3

Query Match 1.5%; Score 10; DB 16; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
Db 113 SSSSSSDSST 122

RESULT 46
US-09-056-019-1
; Sequence 1, Application US/09056019A
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Wizemann, Theresa
; APPLICANT: Masure, H. R.
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Koenig, Scott
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; FILE REFERENCE: 1340-1-017
; CURRENT APPLICATION NUMBER: US/09/056,019A
; CURRENT FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-056-019-1

Query Match 1.5%; Score 10; DB 14; Length 406;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
Db 113 SSSSSSDSST 122

RESULT 47
US-09-287-788-1
; Sequence 1, Application US/09287788B
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Masure, H. R.
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; FILE REFERENCE: 1340-1-017N
; CURRENT APPLICATION NUMBER: US/09/287,788B
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: 60/080,878
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 39

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 406
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-287-788-1

Query Match 1.5%; Score 10; DB 16; Length 406;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
DB 113 SSSSSSDSST 122

RESULT 48
US-09-056-019-24
Sequence 24, Application US/09056019A
GENERAL INFORMATION:
APPLICANT: Tuomanen, Elaine I
APPLICANT: Witzemann, Theresa
APPLICANT: Masure, H. R.
APPLICANT: Johnson, Leslie S.
TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
TITLE OF INVENTION: THEREFROM AND USES THEREOF
FILE REFERENCE: 1340-1-017
CURRENT APPLICATION NUMBER: US/09/056, 019A
CURRENT FILING DATE: 1998-04-07
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 24
LENGTH: 428
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-056-019-24

Query Match 1.5%; Score 10; DB 14; Length 428;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
DB 113 SSSSSSDSST 122

RESULT 49
US-09-287-788-24
Sequence 24, Application US/09287788B
GENERAL INFORMATION:
APPLICANT: Tuomanen, Elaine I
APPLICANT: Masure, H. R.
TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
TITLE OF INVENTION: THEREFROM AND USES THEREOF
FILE REFERENCE: 1340-1-017N
CURRENT APPLICATION NUMBER: US/09/287, 788B
CURRENT FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: 60/080, 878
PRIOR FILING DATE: 1998-04-07
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 24
LENGTH: 428
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-287-788-24

Query Match 1.5%; Score 10; DB 16; Length 428;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
DB 113 SSSSSSDSST 122

RESULT 50
US-09-605-703B-1256
Sequence 1256, Application US/09605703B
GENERAL INFORMATION:
APPLICANT: Pompeius, Markus
APPLICANT: Krieger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberkauer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: BGI-129CP
CURRENT APPLICATION NUMBER: US/09/605, 703B
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: 60/142, 764
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: 60/152, 318
PRIOR FILING DATE: 1999-09-03
NUMBER OF SEQ ID NOS: 2934
SEQ ID NO 1256
LENGTH: 428
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-605-703B-1256

Query Match 1.5%; Score 10; DB 20; Length 428;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
DB 59 SSSSSSDSST 68

Search completed: June 13, 2002, 08:48:47
Job time: 243 sec

• • • • •

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:45:49 ; Search time 20.8 Seconds
(without alignments)
2561.523 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 666
Sequence: 1 KKYDNKNGLIADIGSERRYN.....TQSSNTTPDOQNONPOPAOP 666

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 244551 seqs, 79999522 residues

Word size : 10

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Pending_Patents_AA_New.*

1: /cgn2-6/ptodata/1/paa/PCr_NEW.COMB.pep.*
2: /cgn2-6/ptodata/1/paa/US06_NEW.COMB.pep.*
3: /cgn2-6/ptodata/1/paa/US07_NEW.COMB.pep.*
4: /cgn2-6/ptodata/1/paa/US08_NEW.COMB.pep.*
5: /cgn2-6/ptodata/1/paa/US09_NEW.COMB.pep.*
6: /cgn2-6/ptodata/1/paa/US10_NEW.COMB.pep.*
7: /cgn2-6/ptodata/1/paa/US60_NEW.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	ID	Description
	1	666	100.0	666	4	US-08-961-083-2	Sequence 2, Appli
	2	11	1.7	821	5	US-09-769-744A-122	Sequence 122, App
	3	10	1.5	110	4	US-08-961-083-102	Sequence 102, App
	4	10	1.5	194	6	US-10-105-299-4083	Sequence 4083, Ap
	5	10	1.5	602	7	US-60-360-039-3150	Sequence 3150, Ap
	6	10	1.5	1704	5	US-09-991-262-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-08-961-083-2
; Sequence 2, Application US/08961083

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-961-083-2

Query Match 100.0%; Score 666; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KKYDNKNGLIADIGSERRYNAQNDIPTDLVKAIVSIEDHRRFDHGCIDTIRILGAFLEYN 60
DB	1	KKYDNKNGLIADIGSERRYNAQNDIPTDLVKAIVSIEDHRRFDHGCIDTIRILGAFLEYN 60
QY	61	LOSNSLGGSTLTLOOLIKITFYESTSTSDOTISRAQEAFLAIQLEQKATKOEILFYIYNK 120
DB	61	LOSNSLGGSTLTLOOLIKITFYESTSTSDOTISRAQEAFLAIQLEQKATKOEILFYIYNK 120
QY	121	VYMSNGNYMGQTAQNYGYKDLNNLSLPQLALLAGMPQAPNOYDPYSHPEAODRRNLYL 180
DB	121	VYMSNGNYMGQTAQNYGYKDLNNLSLPQLALLAGMPQAPNOYDPYSHPEAODRRNLYL 180
QY	181	SEKKNNGYISAEQYKAVNTPITDGLQSLKASNTYAVYNDYIKYINOVBEETGYNLLT 240
DB	181	SEKKNNGYISAEQYKAVNTPITDGLQSLKASNTYAVYNDYIKYINOVBEETGYNLLT 240
QY	241	TGMDVYTNVDOEKHLMIDYNTDEYVAYPDELAQVASTIVDSNGKYAQLGARHOSNN 300
DB	241	TGMDVYTNVDOEKHLMIDYNTDEYVAYPDELAQVASTIVDSNGKYAQLGARHOSNN 300
QY	301	VSEGINQAVETNRDWSITMKPTTDYAPALEYGYVDSFATVHDEPNYPCGTNPVYNMNR 360
DB	301	VSEGINQAVETNRDWSITMKPTTDYAPALEYGYVDSFATVHDEPNYPCGTNPVYNMNR 360
QY	361	GYFGNTTLOYALQOSRNVAVETLNKVGUNRAKTFPLNGLGIDYPSIHYSAISSNTTESD 420
DB	361	GYFGNTTLOYALQOSRNVAVETLNKVGUNRAKTFPLNGLGIDYPSIHYSAISSNTTESD 420
QY	421	KRYGASSEKMAAAYAFANGGYTKPMYIHKVYFSGSEKESFNVGTBAMKETTAAMMD 480
DB	421	KRYGASSEKMAAAYAFANGGYTKPMYIHKVYFSGSEKESFNVGTBAMKETTAAMMD 480
QY	481	MMKTYLVYTGGRNAYLAWLPQAGKTGTSNYTDEIEINHIKTSQFVAPDELFACTRYKYSM 540
DB	481	MMKTYLVYTGGRNAYLAWLPQAGKTGTSNYTDEIEINHIKTSQFVAPDELFACTRYKYSM 540
QY	541	AWTGTYSNRLTPLYGNGLTVAAKVYRSMYIYSEGSNPEDWNIPEGLYRNGEFVFNKGAR 600
DB	541	AWTGTYSNRLTPLYGNGLTVAAKVYRSMYIYSEGSNPEDWNIPEGLYRNGEFVFNKGAR 600

QY 601 STWNSAPDPPSTESSSSSDSTSSSTPSTNNSTTTNNNNNTTQOSNTTDDOON 660
 Db 601 STWNSAPDPPSTESSSSSDSTSSSTPSTNNSTTTNNNNNTTQOSNTTDDOON 660
 QY 661 POPAP 666
 Db 661 POPAP 666

RESULT 2

US-09-769-744A-122
 ; Sequence 122, Application US/09769744A
 ; GENERAL INFORMATION:
 ; APPLICANT: Le Page, Richard WF
 ; APPLICANT: Wells, Jeremy M
 ; APPLICANT: Hanniffy, Sean B
 ; APPLICANT: Hansbro, Philip M
 ; TITLE OF INVENTION: Proteins
 ; FILE REFERENCE: PMC/P21122WO
 ; CURRENT APPLICATION NUMBER: US/09/769,744A
 ; CURRENT FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: PCF/GB99/02452
 ; PRIOR FILING DATE: 1999-07-27
 ; PRIOR APPLICATION NUMBER: GB 9816336.3
 ; PRIOR FILING DATE: 1998-07-27
 ; PRIOR APPLICATION NUMBER: US 60/125329
 ; PRIOR FILING DATE: 1999-03-19
 ; NUMBER OF SEQ ID NOS: 196
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 122
 ; LENGTH: 821
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pneumoniae
 US-09-769-744A-122

Query Match 1.7%; Score 11; DB 5; Length 821;
 Best Local Similarity 100.0%; Pred. No. 0.073;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 GGSTLTQOLIK 78
 Db 176 GGSTLTQOLIK 186

RESULT 3

US-08-961-083-102
 ; Sequence 102, Application US/08961083
 ; GENERAL INFORMATION:
 ; APPLICANT: Choi et. al.
 ; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
 ; NUMBER OF SEQUENCES: 452
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/961,083
 ; FILING DATE: 30-Oct-1997
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: <Unknown>
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB340P2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 102:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 110 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 102:
 US-08-961-083-102

Query Match 1.5%; Score 10; DB 4; Length 110;
 Best Local Similarity 100.0%; Pred. No. 0.1;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
 Db 18 SSSSSSDSST 27

RESULT 4

US-10-105-299-4083
 ; Sequence 4083, Application US/10105299
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen, et. al
 ; TITLE OF INVENTION: Human Secreted Proteins
 ; FILE REFERENCE: PS950
 ; CURRENT APPLICATION NUMBER: US/10/105,299
 ; CURRENT FILING DATE: 2002-03-26
 ; NUMBER OF SEQ ID NOS: 15197
 ; Prior Application removed - See File Wrapper or Palm
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4083
 ; LENGTH: 194
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-105-299-4083

Query Match 1.5%; Score 10; DB 6; Length 194;
 Best Local Similarity 100.0%; Pred. No. 0.17;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
 Db 149 SSSSSSDSST 158

RESULT 5

US-60-360-039-3150
 ; Sequence 3150, Application US/60360039
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Chen, Xianfeng
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)A
 ; CURRENT APPLICATION NUMBER: US/60/360,039
 ; CURRENT FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 3150
 ; LENGTH: 602
 ; TYPE: PRT
 ; ORGANISM: Neurospora crassa
 ; FEATURE:

NAME/KEY: unsure
LOCATION: (1)...(602)
OTHER INFORMATION: unsure at all Xaa locations
US-60-360-039-3150

Query Match 1.5%; Score 10; DB 7; Length 602;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 615 ESSSSSDSS 624
Db 311 ESSSSSDSS 320

RESULT 6
US-09-991-262-40
Sequence 40, Application US/09991262
GENERAL INFORMATION:
APPLICANT: Christlan, P. D., Gordon, K. H.J., Hanzlik, T. N.
TITLE OF INVENTION: Insect Viruses and Their Uses in
Protecting Plants
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr Hobach Test Albritton & Herbert LLP
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/991,262
FILING DATE: 20-NOV-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/234,238
FILING DATE: 20-JAN-1999
APPLICATION NUMBER: US 08/485,355
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/440,522
FILING DATE: 12-MAY-1995
APPLICATION NUMBER: US 08/089,372
FILING DATE: 08-JUL-1993
APPLICATION NUMBER: AU PL4081/92
FILING DATE: 14-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Treccartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-58631-3/RFT/DSS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-09-991-262-40

Query Match 1.5%; Score 10; DB 5; Length 1704;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 SSSSSSDSTS 626

Db 1452 SSSSSSDSTS 1461

Search completed: June 13, 2002, 08:49:37
Job time: 228 sec

08/961083

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:40:34 ; Search time 34.61 Seconds
(without alignments)
2137.392 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484
Sequence: 1 KIDNKNQIADLGSERRVN.....TQSNTPDQONQNPQAP 666

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802:.*
1: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3484	100.0	666	19	AAW55063
2	3466	99.5	682	17	AAW04359
3	3466	99.5	719	22	AAU37830
4	2208	63.4	420	20	AAV56106
5	1751	50.3	778	22	AAU35135
6	1621	46.5	320	16	AAU70153
7	1361	39.1	462	22	AAU03646
8	817.5	23.5	727	22	AAU34286
9	817.5	23.5	727	22	AAU37184
10	640	18.4	805	21	AAV75603
11	638	18.3	805	21	AAV75602

12	636	18.3	731	19	AAW44849	S. pneumoniae peni
13	635.5	18.2	805	21	AAV75601	Neisseria gonorrhoe
14	625	17.9	660	18	AAW55545	H. pylori ORF 06ep
15	625	17.9	660	18	AAW55576	H. pylori ORF 06ep
16	608.5	17.5	596	20	AAW89836	Protein encoded by
17	586	16.8	812	22	AAU78604	Lawsonia intracell
18	567	16.3	828	17	AAW04357	E. coli penicillin
19	555	15.9	850	17	AAW86955	E. coli PBP 1A tra
20	552	15.8	774	22	AAU36453	Pseudomonas aerugi
21	539.5	15.5	536	20	AAW89885	Antigen 2 from clu
22	529.5	15.2	821	21	AAV81757	Streptococcus pneu
23	326	15.1	823	17	AAW04358	E. coli penicillin
24	526	15.1	844	22	AAU34446	E. coli cellular p
25	525.5	15.1	781	22	AAU35684	Haemophilus influe
26	519.5	14.9	846	22	AAU38311	Salmonella typhi c
27	514	14.8	844	17	AAW86953	E. coli PBP 1B tra
28	512	14.7	844	17	AAW86953	E. coli PBP 1B tra
29	508	14.6	833	22	AAU55400	Propionibacterium
30	489	14.0	836	17	AAW86954	E. coli PBP 1B tra
31	457	13.1	447	18	AAW20510	H. pylori inner me
32	457	13.1	799	22	AAU49462	Propionibacterium
33	450.5	12.9	770	22	AAU94353	Escherichia coli p
34	441.5	12.7	795	22	AAU80056	C glutamicum prote
35	437.5	12.6	678	22	AAU81107	Mycobacterium tube
36	422	12.1	810	22	AAU33052	C glutamicum prote
37	419	12.0	822	22	AAU81232	Mycobacterium tube
38	377	10.8	344	18	AAW55362	H. pylori ORF 14gp
39	377	10.8	344	18	AAW20917	H. pylori cell env
40	370.5	10.6	532	17	AAW86957	E. coli penicillin
41	370.5	10.6	553	17	AAW86956	E. coli penicillin
42	341.5	9.8	231	20	AAW0230	R. eutropha Mgt pa
43	288.5	8.3	269	19	AAW61395	Staphylococcus aur
44	287	8.2	338	22	ABG24982	Novel human diagno
45	280.5	8.1	269	22	AAU81893	S. epidermidis ope

ALIGNMENTS

RESULT 1	
AAW55063	standard; Protein: 666 AA.
XX	
AC	AAW55063;
XX	
DF	02-OCT-1998 (first entry)
XX	
DE	Streptococcus pneumoniae SP001 protein.
XX	
KW	Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
KW	detection; pneumonia; otitis media; meningitis.
XX	
OS	Streptococcus pneumoniae.
XX	
PN	MO9818930-A2.
XX	
PD	07-MAY-1998.
XX	
PF	30-OCT-1997; 97WO-US19422.
XX	
PR	31-OCT-1996; 96US-0029960.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Choi GH, Hromocky A, Johnson LS, Kunsch CA;
XX	
DR	WPI: 1998-272224/24.
XX	
DR	N-PSDB; AAV27323.
XX	
PT	Nucleic acid encoding antigenic peptide(s) from Streptococcus
PT	pneumoniae - or their epitope-containing fragments, useful in
XX	protective or therapeutic vaccines, and for diagnosis

PS Claim 11; Page 48; 118pp; English.

CC The present sequence represents a protein from *Streptococcus pneumoniae*.
 CC The nucleic acid sequence encoding the *Streptococcus pneumoniae* protein
 CC can be useful in vaccines for inducing protective antibodies against
 CC *Streptococcus pneumoniae*, for treatment or prevention of infection e.g.
 CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid
 CC are used to detect *Streptococcus pneumoniae* infection (by usual hybridisation or
 CC amplification methods), also for isolating *Streptococcus pneumoniae* or their
 CC allelic variants. The protein can be used similarly to detect specific
 CC antibodies in standard immunoassays, especially for diagnosing or
 CC monitoring infections. Antibodies which bind the protein are used to
 CC detect corresponding antigens, to purify the protein and for passive
 CC immunisation (optionally coupled to a toxin). Vaccines are administered,
 CC e.g. by injection, orally or through the skin, typically at 0.01-1000
 CC (especially 10-300) μ g/ml per dose.

CC Sequence 666 AA:

Query Match 100.0%; Score 3484; DB 19; Length 666;
 Best Local Similarity 100.0%; Pred. No. 1,1e-234;
 Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KIYDNKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLENN 60
 DB 1 KIYDNKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLENN 60
 OY 61 LOSNSLQGGSTLTQOLIKLTYFSTSDQTSIRKQAEAWLALQLEKATKQELITYYINK 120
 DB 61 LQSNLSLQGGSTLTQOLIKLTYFSTSDQTSIRKQAEAWLALQLEKATKQELITYYINK 120
 OY 121 VYMSNGNYGMQTAQNYKYKDLNNLSLPOLALLAGMPQAPNODPYSHPEAADRNLVYL 180
 DB 121 VYMSNGNYGMQTAQNYKYKDLNNLSLPOLALLAGMPQAPNODPYSHPEAADRNLVYL 180
 OY 181 SEMKNOGYTSABQYKAVNPTITDGLQSLKSNPYAYMDNVLKEVINOVEEETGYNLLT 240
 DB 181 SEMKNOGYTSABQYKAVNPTITDGLQSLKSNPYAYMDNVLKEVINOVEEETGYNLLT 240
 OY 241 TGMDDVYTNDOERAKHLMQITMDEVVAAPDDELQVASTIVVSNCKVIAQAGARRQSSN 300
 DB 241 TGMDDVYTNDOERAKHLMQITMDEVVAAPDDELQVASTIVVSNCKVIAQAGARRQSSN 300
 OY 301 VSEFGINQAEVETNRDMGSTMKPTTDYAPALEYGYDSTATIVHDEPNYPGTTPVYNNMDR 360
 DB 301 VSEFGINQAEVETNRDMGSTMKPTTDYAPALEYGYDSTATIVHDEPNYPGTTPVYNNMDR 360
 OY 361 GFYGNITLQYALQOSRNPVAVETLNVKGLNRAKTFINGLIGIDYPSIHYSNALISSNTTESD 420
 DB 361 GFYGNITLQYALQOSRNPVAVETLNVKGLNRAKTFINGLIGIDYPSIHYSNALISSNTTESD 420
 OY 421 KRYGASSEKMAAAYAFAGNGTYVYKPMYTHKVFVSDGSEKESNVTGRAMKETTAYMMTD 480
 DB 421 KRYGASSEKMAAAYAFAGNGTYVYKPMYTHKVFVSDGSEKESNVTGRAMKETTAYMMTD 480
 OY 481 NMKTYLTYGTGRNAYLAWLPQAGKTGTSNYTDEIEENHKTQSFVYPPDGLFAGYTRKRYGM 540
 DB 481 NMKTYLTYGTGRNAYLAWLPQAGKTGTSNYTDEIEENHKTQSFVYPPDGLFAGYTRKRYGM 540
 OY 541 AVMTGYSNRLPLVNGSLTVAAYKVRSMATVYLSSEGSNPEDMNIPEGLYRNGEVFENGAR 600
 DB 541 AVMTGYSNRLPLVNGSLTVAAYKVRSMATVYLSSEGSNPEDMNIPEGLYRNGEVFENGAR 600
 OY 601 STWNSPAPQPPSTESSSSSSSSSTPSTNNSTTTTPNNNTTQOQSNTPPDQONON 660
 DB 601 STWNSPAPQPPSTESSSSSSSSSTPSTNNSTTTTPNNNTTQOQSNTPPDQONON 660
 OY 661 POPAOP 666
 DB 661 POPAOP 666

RESULT 2

ID AAW04359 standard; Protein; 682 AA.

AAW04359;

03-DEC-1996 (first entry)

S. pneumoniae penicillin binding protein 1A soluble variant.

Penicillin binding protein; PBP 1A; bifunctional protein;

transglycosylase; transpeptidase; identification; assay; inhibitor;

antibiotic resistant; bacteria; soluble variant; protein structure;

X-ray crystallography; determination.

Streptococcus pneumoniae.

GB2290792-A.

10-JAN-1996.

29-JUN-1995; 95GB-0013306.

24-NOV-1994; 94SE-0004072.

01-JUL-1994; 94TN-0000580.

(ASTR) ASTRA AB.

Balganesh TS, Town CM;

WPI; 1996-042332/05.

N-PSDB; AAT08027.

Sol. derivs. of bifunctional penicillin binding protein (BPBP) -

opt. lack transglycosylase activity, useful to identify and assay

for antibodies or cpds. which bind BPBPs

Claim 2; Pages 65-68; 108pp; English.

The present sequence is a soluble variant of the S. pneumoniae

penicillin binding protein (PBP) 1A, which lacks the 1st 38 amino

acid residues of the wild type protein. Wild type PBP is a

bifunctional protein, which binds the cell membrane when expressed

in a bacterial cell, having transglycosylase and transpeptidase

activities. The variant protein (NCIMB 40665) in conjunction with

a labelled anti-bifunctional PBP monoclonal antibody, can be used

to identify and assay for cpds. which bind bifunctional PBP. Such

cpds., as inhibitors of bifunctional PBP have a potential use in

therapeutic cpds. which inhibit the growth of antibiotic resistant

bacteria. The soluble variant may also be used in X-ray

crystallography.

Sequence 682 AA:

Query Match 99.5%; Score 3466; DB 17; Length 682;
 Best Local Similarity 99.4%; Pred. No. 2e-233;
 Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KIYDNKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLENN 60
 DB 1 KIYDNKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLENN 60
 OY 61 LOSNSLQGGSTLTQOLIKLTYFSTSDQTSIRKQAEAWLALQLEKATKQELITYYINK 120
 DB 61 LOSNSLQGGSTLTQOLIKLTYFSTSDQTSIRKQAEAWLALQLEKATKQELITYYINK 120
 OY 121 VYMSNGNYGMQTAQNYKYKDLNNLSLPOLALLAGMPQAPNODPYSHPEAADRNLVYL 180
 DB 121 VYMSNGNYGMQTAQNYKYKDLNNLSLPOLALLAGMPQAPNODPYSHPEAADRNLVYL 180
 OY 181 SEMKNOGYTSABQYKAVNPTITDGLQSLKSNPYAYMDNVLKEVINOVEEETGYNLLT 240
 DB 181 SEMKNOGYTSABQYKAVNPTITDGLQSLKSNPYAYMDNVLKEVINOVEEETGYNLLT 240

Db 197 semknggyisaeyekavntpitdglqslksasnypaymdnlykevinqveetgynllt 256
QY 241 TGMVDYTNVQBAQKHLMDIYNFTDEYVAYPDDDELQVASTIVDVSNKGKIAQLGARHQSNN 300
Db 257 tgmdivtnvdqeaqkhlwdiynftdeyaypdddelqvastivdvsnkgkviaqlgarhqsnn 316
QY 301 VSFGINQAVETNRDWSGTMKPIITDYAPALEYGVYDSTATIVHDEPNYPGTNTPVYNWDR 360
Db 317 vsfginqavetnrdwsgtmkpiitdyapalegyvystativhdepyngtntpynwdr 376
QY 361 GYFGNITLQYALQOQSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHYSNALSSNTTSED 420
Db 377 gyfgnitiqyalqqsgrnvpavetlnkvlnraktfnglgiidypsihynalsnttssed 436
QY 421 KKYGASSEKMAAAYAAFAANGGTYYPKMYIHKVVFSDGSEKFSNVGTRAMKETTAYMMTD 480
Db 437 kkygassekmaaaayaaafanggttyypkmyihkvvfsgdsekefsnvgtramkettaymtd 496
QY 481 MMKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEEENHIKTSQFVAPDELFAGYTRKYSM 540
Db 497 mmktvlsygtgrnaylawlpqagktgtsnytddeenhihtskfsqfvpdelafagytrkysm 556
QY 541 AWTGYSNRLPLVGNGLTVAAKVYRSMYTLSEGSNPEDWNIPGELYNGEFVFKNGAR 600
Db 557 awtgyisnrlplvnglntvaakvyrsmttylsegnsnpedwnipeglyrngefvfkngar 616
QY 601 STWNSPAPQPPSTESSSSSSSTSSQSSSTPSTNNSTTTNPNNNTQOSNTTTPQOQON 660
Db 617 stwspapqppstessssssstssqsssttpstnnstttnpnnntqosnttppdqon 676
QY 661 PQAQAP 666
Db 677 pqapqp 682
RESULT 3
AAU37830
ID AAU37830 standard; Protein: 719 AA.
AC AAU37830;
XX
DT 14-FEB-2002 (first entry)
DE Streptococcus pneumoniae cellular proliferation protein #259.
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Streptococcus pneumoniae.
XX WO200170955-A2.
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206849P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlson KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI: 2001-611495/70.
DR N-PSDB: AAS55689.
XX
PT New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 13423; 511pp; English.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 719 AA;
Query Match 99.5%; Score 3466; DB 22; Length 719;
Best Local Similarity 99.4%; Pred. No. 2.1e-233;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 KIYDNKNQIADLGSERRVNAOANDIPTDLVKAIVSIEDHRRFDHGRIDITRILGAFLRN 60
Db 54 kiydnnkqliadlgseerrvnaandiptdlvkaivsiedhrfrfdhgriditirilgaflrn 113
QY 61 LOSNSLQGGSTITQOLIKTYFSTSDQTSIRKQAQEWLAIQLEOKATYKQBELTYIYINK 120
Db 114 lqsnslqggstaltqliklityfstsdqtsirkqaqewlaiaqlqekatkqeltyyink 173
QY 121 VYMSNGNYGMQTAQNYGKDLNLSLPOLALLAGNPQAPNOYDPSHPEAQRNLVL 180
Db 174 vymsngnygmtaagnyygkdlnlsipqlallagmpqapndypshpeaqdrnlvl 233
QY 181 SEMKNGYISAEQYKAVNTPITDGLQSLKSASNPAYMDNLYKEVINQVEETGYNLLT 240
Db 234 semknggyisaeyekavntpitdglqslksasnypaymdnlykevinqveetgynllt 293
QY 241 TGMVDYTNVQBAQKHLMDIYNFTDEYVAYPDDDELQVASTIVDVSNKGKIAQLGARHQSNN 300
Db 294 tgmdivtnvdqeaqkhlwdiynftdeyaypdddelqvastivdvsnkgkviaqlgarhqsnn 353
QY 301 VSFGINQAVETNRDWSGTMKPIITDYAPALEYGVYDSTATIVHDEPNYPGTNTPVYNWDR 360
Db 354 vsfginqavetnrdwsgtmkpiitdyapalegyvystativhdepyngtntpynwdr 413
QY 361 GYFGNITLQYALQOQSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHYSNALSSNTTSED 420
Db 414 gyfgnitiqyalqqsgrnvpavetlnkvlnraktfnglgiidypsihynalsnttssed 473
QY 421 KKYGASSEKMAAAYAAFAANGGTYYPKMYIHKVVFSDGSEKFSNVGTRAMKETTAYMMTD 480
Db 474 kkygassekmaaaayaaafanggttyypkmyihkvvfsgdsekefsnvgtramkettaymtd 533
QY 481 MMKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEEENHIKTSQFVAPDELFAGYTRKYSM 540
Db 534 mmktvlsygtgrnaylawlpqagktgtsnytddeenhihtskfsqfvpdelafagytrkysm 593
QY 541 AWTGYSNRLPLVGNGLTVAAKVYRSMYTLSEGSNPEDWNIPGELYNGEFVFKNGAR 600
Db 594 awtgyisnrlplvnglntvaakvyrsmttylsegnsnpedwnipeglyrngefvfkngar 653
QY 601 STWNSPAPQPPSTESSSSSSSTSSQSSSTPSTNNSTTTNPNNNTQOSNTTTPQOQON 660
Db 653 stwspapqppstessssssstssqsssttpstnnstttnpnnntqosnttppdqon 660


```

Db 245 tgmdivytnvdeaqkhlwdlyntdeyaypddelqvastivdsvngkviaglgarhgsn 304
QY 301 VSFQINQAVETNRDWG 316
Db 305 vsfgingqvavetnrdwg 320

RESULT 7
ID AAU03646
XX AAU03646 standard; Protein; 462 AA.
AC AAU03646;
DT 12-SEP-2001 (first entry)
XX Group B Streptococcus antigenic protein, ID-122.
XX Group B Streptococcus; encapsulated bacterium; therapeutic; sepsis;
KW meningitis; neonate; antigenic; vaccine; infection; genital tract;
KW capsid polysaccharide vaccination.
XX Streptococcus agalactiae.
OS
XX WO200132882-A2.
XX 10-MAY-2001.
XX 07-SEP-2000; 2000WO-GB03437.
XX 07-SEP-1999; 99GB-0021125.
XX (MICR-) MICROBIAL TECHNIKS LTD.
XX Le Page RWF, Wells JM, Hanniffy SB;
XX WPI; 2001-316444/33.
XX N-PSDB; AAS07063.
XX New polypeptides derived from Streptococcus agalactiae are useful to
PT provide detection of, and vaccination against, Group B Streptococcus
PT infections, particularly to prevent infection in neonates -
XX Claim 1; Fig 1; 178pp; English.
XX AAU03601-AAU03722 represent Group B Streptococcus (Streptococcus
CC agalactiae) amino acid sequences of the invention.. S. agalactiae is an
CC encapsulated bacterium which is a major pathogen of humans causing sepsis
CC and meningitis in neonates as well as adults. The S. agalactiae antigenic
CC polypeptides are used to vaccinate against Group B Streptococcus
CC infections, particularly to prevent infection in new born children
CC arising from the maternal genital tract. An immunogenic composition is
CC useful in the preparation of a medicament for the treatment or
CC prophylaxis of Group B Streptococcus infection. The invention does not
CC have the disadvantages of varied response rate associated with prior art
CC capsid polysaccharide vaccination against Group B Streptococcus.
XX Sequence 462 AA;

Query Match 39.1%; Score 1361; DB 22; Length 462;
Best Local Similarity 62.0%; Pred. No. 1e-86;
Matches 251; Conservative 73; Mismatches 81; Indels 0; Gaps 0;
QY 2 IYDNKNQLIADLGSERRVNAQNDIPTDLKAIYSDHREDFHRTGIDTIRILGAFLRNL 61
Db 58 vydgunkliadlgseksvsadslpnlvnaitscedkfrfkhrgvdiylilgaahnl 117
QY 62 QSNLSQGGSTUQOLIKLTYFSTSTQTSIRKQAEAWLAQLBQKATKQBIILYINKV 121
Db 118 vssntqggstldqklklayfstnksqtkrksqewlqmrkykeilfyinkv 177
QY 122 YMSNGNYGMTAAQNYGKDLNLSLPQALLAGMPQAPNOYDPSHPEAAQDRRLVLS 181

```

```

Db 178 ymgngngmrttaksyfgkdkelsiaqlallagipaqpydyknpesaqrntvqlq 237
QY 182 EMKNQGYISAEQYKAVNTPTDGLQSLKASNSPAYMDNYLKEVINQVEEETGYNLLTT 241
Db 238 qmyqdkniskkeydqavatpvtgkkelkqkstyypkymdnylkqvisevkgktgkdifta 297
QY 242 GMDVYTNVDOEAOKHLWDIYNTDEYVAYPDDELQVASTIVDSNGKVIAGLGAHQSSNV 301
Db 298 glkvytnintdaqkqlydiynsdtyiaypnnelqiaastmdatngkviaglggrhqneni 357
QY 302 SFGINQAVETNRDWGSTMKPTIDYAPALEYGVYDSTATIVHDEPNYPGTPTVYNDRG 361
Db 358 sfgtngsvlttdrgwgtmkpisyapaldsgvynustggsindsvyypgtetqlydwdrq 417
QY 362 YFGNITLQYALQQSRNVPAYETLNKVLNRAKTEFLNGLIDYPSI 406
Db 418 ymgwmsmqtaiqgrnvpavraleaagldaeksflekligyypem 462

RESULT 8
ID AAU34286
XX AAU34286 standard; Protein; 727 AA.
AC AAU34286;
DT 14-FEB-2002 (first entry)
XX Staphylococcus aureus cellular proliferation protein #562.
DE Staphylococcus aureus cellular proliferation protein #562.
XX Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX Staphylococcus aureus.
XX WO200170955-A2.
XX 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US09180.
XX 21-MAR-2000; 2000US-191078P.
XX 23-MAY-2000; 2000US-206848P.
XX 26-MAY-2000; 2000US-207727P.
XX 23-OCT-2000; 2000US-242578P.
XX 27-NOV-2000; 2000US-253625P.
XX 22-DEC-2000; 2000US-257931P.
XX 16-FEB-2001; 2001US-269308P.
XX (ELIT-) ELITRA PHARM INC.
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
XX N-PSDB; AAS52145.
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
PT
XX Example 3; Seq ID No 5782; 511pp; English.
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery

```


CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ

Sequence 727 AA;

Query Match 23.5%; Score 817.5; DB 22; Length 727;
Best Local Similarity 32.1%; Pred. No. 1.8e-48;
Matches 222; Conservative 128; Mismatches 261; Indels 81; Gaps 22;

QY 1 KIYDNKNQIADL-GSERVNAQANDIPTDLVKAIVSIEDHRRFFDHRGIDTIRILGAFRL 59
DB 76 KIYDNKgelvktldngqrhehvnldkvpksmkdaviatdnrfyehgaIdykrIfgaIgk 135
QY 60 NLOSN-SLOGGSTLTQQLIKLYFTSTSDQTSIRKAQEAWLAIQLEQKATQOEILTYI 118
DB 136 nltgfgfsegastltgqvvdafsl---qhksigrkaeaylsyrleqyskddifqyvl 192
QY 119 NKVYMSNGYGMOTRAQNYGKDLNNLSLPQALLACMPQAPNOYDPYSHPEAQRRL 178
DB 193 nkIyysdgvgtgikaakyfynkldlnlaeeaylaglpqvpnnynIydhpkaaedrkt 252
QY 179 VLSEMKNGYISAEQYKA-----VN-TPITDGLQLSKASNPAYMDNYLKEVIN 228
DB 253 vlylmhYkrlcdkWedakkldkanlvnrCpeqrnidtnqdsyfnvfkseImn 312
QY 229 Q--VEEETGYNLLTTCMDYTNVDOEAQKHLWDIYNTDEYVAYPDEQLQVASTIVDVSNG 286
DB 313 nkafkdenlgvlgsgikIytnmdkdvktlqn--dvdngsfyknkdgqvgatIdsktg 370
QY 287 KVIAQLGARHOSNNVFGINQAVETNRDWMGTMKPIIDYAPALEYGVVDSTATIVHDEPY 346
DB 371 glvaIsgrdrfdkvvnn--rnqatdph-ptgssalkpflaygpaiemkmkwnhaiqdesy 427
QY 347 NYPGNTPTVYNWDRGYFNITLQYALQOSRNPVAVETLNKV---GLNRAKTFNLCLGID 402
DB 428 qvdgst--frnydtkshgtcvslYdairgsfnlpakawgsvkqngnagdapkkfaaklgIn 485
QY 403 YP-SIHYNNAISNPTSDKKYGSASERKMAAAYAFANGTGYKPMYIHKVVFSDGSEKE 461
DB 486 yegdIgpsevlggsase-----fsptqlasaafaiaanggtynnahsqkvvtrdgetie 539
QY 462 FSNVGTAMKETAYMMTDMK-TVLTYGTGRNAYLAMPQAGKGTGTSNYTDEEIHNIK 520
DB 540 ydhtshkamsdytYmlaemlkgtfkpgysayghvgvnmgakgtgtygaety----- 594
QY 521 TSQFVAPDEL-----FAGYTRKYSMAVWTGYSNRLTPLVGNGLTVAAKVYRSMMYLSG 575
DB 595 -sqynlpdnaaakdvwingtptqytnsvvmgtsk-----vkqy-----g 631
QY 576 SNP-----EDWNIPEGLYRNGEFVFKNGARSTWNSPAPQOPQSTESSSSDSTSQSSST 631
DB 632 ensfghsqgqeypqflyen--vmskissr---dgedfkrpssvsgsipsinvgsgqdnt 686
QY 632 TPST---NNSTTNPNNNTQSNNTTPDQONQ 660
DB 687 tnrschggsdtsanssgtaqsnntsrsgsrn 718

RESULT 9

AAU37184

ID AAU37184 standard; Protein; 727 AA.

XX

AC AAU37184;

XX

DT 14-FEB-2002 (first entry)

XX

DE Staphylococcus aureus cellular proliferation protein #1354.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR N-PSDB; AAS55043.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 12777; 511pp; English.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 727 AA;

Query Match 23.5%; Score 817.5; DB 22; Length 727;
Best Local Similarity 32.1%; Pred. No. 1.8e-48;
Matches 222; Conservative 128; Mismatches 261; Indels 81; Gaps 22;

QY 1 KIYDNKNQIADL-GSERVNAQANDIPTDLVKAIVSIEDHRRFFDHRGIDTIRILGAFRL 59
DB 76 KIYDNKgelvktldngqrhehvnldkvpksmkdaviatdnrfyehgaIdykrIfgaIgk 135
QY 60 NLOSN-SLOGGSTLTQQLIKLYFTSTSDQTSIRKAQEAWLAIQLEQKATQOEILTYI 118
DB 136 nltgfgfsegastltgqvvdafsl---qhksigrkaeaylsyrleqyskddifqyvl 192
QY 119 NKVYMSNGYGMOTRAQNYGKDLNNLSLPQALLACMPQAPNOYDPYSHPEAQRRL 178
DB 193 nkIyysdgvgtgikaakyfynkldlnlaeeaylaglpqvpnnynIydhpkaaedrkt 252

trans-peptidation activity; infection; therapy.

Streptococcus pneumoniae.

EP837132-A2.

22-APR-1998.

17-OCT-1997: 97EP-0308288.

17-OCT-1996: 96US-0731716:

(ELIL) LILLY & CO ELI.

Hoskins JA, Jaskunas SR, Norris FH, Rockey PK, Rosteck PR;

Zhao G;

WPI; 1998-219114/20.

N-PSDB; AAV19373.

Recombinant *Streptococcus pneumoniae* penicillin-binding protein - useful in screening assays for antibacterial agents

Claim 1; Page 19-21; 28pp; English.

This sequence is the *Streptococcus pneumoniae* penicillin-binding protein, designated PBP-Nv, of the invention. The protein is useful in screening assays for compounds that bind to PBP-Nv or inhibit the transglycosylase or trans-peptidation activity of PBP-Nv. Such compounds would be useful as antibacterial agents for treating *S. pneumoniae* infections.

Sequence 731 AA;

Query Match 18.3%; Score 636; DB 19; Length 731;

Best Local Similarity 27.78; Pred. No. 8.5e-36;

Matches 192; Conservative 108; Mismatches 252; Indels 142; Gaps 19;

QY 2 IYDNKNQLIADLGSERRVNAQANDIPTDLVKAIVSIEDHRFFDHRGIDTIRILGAFLRNL 61

[illegible]

62 QNSLOGGSTLT00LIKLTTFSTSTSDOTISRKAQEAWLAIQLEQKATKQEILTYINKV 121

[illegible]

07 122 YMSNGNYGMOTAAONVYGKDI NNI.SI.POI.AI.I.AGMPDAPNOYDPPYSHPEAAODRRNT.VI.S 181

QY 122 YMSNGNIGMUTAAQNYIGADLNNLSLFQLALLAGMPQAFNQIDPISHPEAAQDQRNVLVS 101

100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098

QY 182 EMKNQGYISAEQYEKAVNTPIT----DGLQSLKSASNYPAYMDNYLKVEI---NQVEEET 234

Db 328 ---ivnnqyriytdqnyqanmqiventytslfpraedgtfaqsgsvaleptkgvrgvv 384

QY 293 GARHQSSNVSF-GINQAVEINKRDWGSIMKRFIPDIAFALCEIGVIDSIATTVHDEFI 340

nb 385 aavadndktafrnfvataskrspastikp|vyvtpaveagwalnkaldnhtmaydsvkv 444

QY 347 -NYPGTNT----PVYNWDRGYFGNITLQYALQQSRNVPAVEITLKNKVG LNRKATFLNGLGI 401

pb 115 dñvæqjktærwmmv-----s[æes[n]navatvnd]avdkæ----- 482

QY 402 DYPYSHYNAISSNTTESDKKYGASSEK-----MAAAYAAAFANGGT 442

[illegible]

QY 443 YYKPMYIHKVVFSDGSE-KEFSNVGTRAMKETTAYMMTDMKTVLTYGTGRNAYLAWLPQ 501

Db 165 nqtffghygyvktaslygfkpkldkltlkeitmvalprapsfydptknlfsrand 224

Qy 179 VLSEKMGYISAEQYKAVN--TPITDGLQSLKSASNPAYMDNLYKEVINQVEEETGYN 237

Db 225 ilrrlyslgwissnelksalnepivynqtstgniapy-----vdevlklqldlg-- 276

Qy 238 LTTGMDVYTNVQEAQKHLWDIY-----NTDEVVAYPDDELQVASTIV 281

Db 277 lktggytkltdldyqrlaleslrfghqkilekiaekpktnasndkdndlnasmivt 336

Qy 282 DVSNKGVIQAQCARHQSNNVSGFNGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIV 341

Db 337 ectsgkialvggidykks---afnratqakrqfsgaikipfv-yqlafdnngy--sttski 390

Qy 342 HDEPNYPTGNTTPVY-----NWDGFGNITLQYALQOOSRNPVAVETLNKVG 388

Db 391 pdtarnfengn---ysknsqvphawhpsnytrkfglvtlqlealslnlatinsdqlg 447

Qy 389 LNRKATFLNGLIDYPSIHYSHNAISSNTTESDKKYGASSEKMAAAYAAFANGGTYYPKMY 448

Db 448 fekiyqslsdmgfknlpkdlslvigs-----faispidaaeakyslfsnygtmalkpml 499

Qy 449 IHKVVFSGSEKESNVGCTRAM--KETTAYMMTDMKTVLTGTGRNAYLAWLPOAGKTGT 507

Db 500 iesitnqqnevktftptietkkitkeqaflltsalmdavengtgsarikgleiagktgt 559

Qy 508 SNYDEEIEENHIKTSQFVAPDELFGYTRKYSMAVWTGYSNRLTPLYVGNGLT---VAAKV 564

Db 603 ysyfmrnilaiepslkrfdvpgkgrk--eivdklipyyspnsitp--tpkktdds 654

RESULT 15

AAW55576

ID AAW55576 standard; Protein: 660 AA.

XX AC AAW55576;

XX AC AAW55576;

DT 07-JUL-1998 (first entry)

XX DE H. pylori ORF 06p30223_4698838_f2_55 protein.

XX KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;

XX KW identification; binding compound; bacteria; life cycle; activator;

XX KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.

XX OS Helicobacter pylori.

XX PN W09737044-A1.

XX PD 09-OCT-1997.

XX PF 27-MAR-1997; 97WO-US05223.

XX PR 06-DEC-1996; 96US-0761318.

XX PR 29-APR-1996; 96US-0625811.

XX PR 02-APR-1996; 96US-0758731.

XX PR 25-OCT-1996; 96US-0736905.

XX PR 28-OCT-1996; 96US-0738859.

XX PA (ASTR) ASTRA AB.

XX PI Alm RA, Smith D;

XX XX

DR WPI; 1997-503122/46.

DR N-PSDB; AAW24985.

XX XX

PT Helicobacter pylori nucleic acid sequences and encoded

PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori

PT infection and for diagnosis of H. pylori infection

XX Claim 14; Pages 780-781; 1145pp; English.

XX This sequence is a H. pylori protein of unspecified function.

CC The protein may be used in a vaccine to prevent or treat H. pylori

CC infection or to identify H. pylori polypeptide binding compounds,

CC useful as potential H. pylori life cycle activators or inhibitors. The

CC DNA and probes derived from it may be used for the identification of

CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic

CC acid sequences complementary to the DNA act as antisense sequences and

CC can be used to prevent the translation of H. pylori mRNA. Antibodies

CC against the protein can be used in immunoassays to evaluate the abundance

CC and distribution of H. pylori-specific antigens. The genomic sequence of

CC H. pylori (ATCC 35679) was determined from overlapping contigs generated

CC by mechanically shearing the bacterial DNA. The sequences were analysed

CC for ORF of at least 180 nucleotides, and the predicted coding regions

CC defined by computer evaluation. To identify likely H. pylori antigens for

CC vaccine development, the amino acid sequences predicted from various ORF

CC were analysed for significant homology to other known or exported

CC membrane proteins. Having identified and determined the sequences of

CC interest, particular regions can be isolated from H. pylori by PCR

CC amplification for recombinant polypeptide production, e.g. in E. coli

CC hosts.

XX XX Sequence 660 AA;

Qy Query Match 17.9%; Score 625; DB 18; Length 660;

Best Local Similarity 28.2%; Pred. No. 4.3e-35;

Matches 185; Conservative 129; Mismatches 255; Indels 86; Gaps 22;

Qy 1 KIYDNKNOLIADL-GSERVRNAQANDIPTDLVKAIVSIEDHFRHGRIDITRIILGAFELR 59

Db 48 qildkrgrlianiydkerfyarfeeiiprfvesllavedtlfheggjindavnamrak 107

Qy 60 NLOSNS-LQGGSTLTQOLIKLTYFSTSTSDOTISKQAEAWLAIOLEKATKQEILTYYI 118

Db 108 naksgryteggstltqqlvknmvl---trektitrlkealisiriekvlseelleryi 164

Qy 119 NKVYMSNGYMGQTAQNYGKDLNNLSLPQLALLAGPQAPNOYDPYSHPEAAODRRNL 178

Db 165 nqtffghygyvktaslygfkpkldkltlkeitmvalprapsfydptknlfsrand 224

Qy 179 VLSEKMGYISAEQYKAVN--TPITDGLQSLKSASNPAYMDNLYKEVINQVEEETGYN 237

Db 225 ilrrlyslgwissnelksalnepivynqtstgniapy-----vdevlklqldlg-- 276

Qy 238 LTTGMDVYTNVQEAQKHLWDIY-----NTDEVVAYPDDELQVASTIV 281

Db 277 lktggytkltdldyqrlaleslrfghqkilekiaekpktnasndkdndlnasmivt 336

Qy 282 DVSNKGVIQAQCARHQSNNVSGFNGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIV 341

Db 337 ectsgkialvggidykks---afnratqakrqfsgaikipfv-yqlafdnngy--sttski 390

Qy 342 HDEPNYPTGNTTPVY-----NWDGFGNITLQYALQOOSRNPVAVETLNKVG 388

Db 391 pdtarnfengn---ysknsqvphawhpsnytrkfglvtlqlealslnlatinsdqlg 447

Qy 389 LNRKATFLNGLIDYPSIHYSHNAISSNTTESDKKYGASSEKMAAAYAAFANGGTYYPKMY 448

Db 448 fekiyqslsdmgfknlpkdlslvigs-----faispidaaeakyslfsnygtmalkpml 499

Qy 449 IHKVVFSGSEKESNVGCTRAM--KETTAYMMTDMKTVLTGTGRNAYLAWLPOAGKTGT 507

Db 500 iesitnqqnevktftptietkkitkeqaflltsalmdavengtgsarikgleiagktgt 559

Qy 508 SNYDEEIEENHIKTSQFVAPDELFGYTRKYSMAVWTGYSNRLTPLYVGNGLT---VAAKV 564

Db 560 sn-----nni-----dawfigftptlqslvifgrddn--tp-igkgatggvvsapv 602

Qy 565 YRSMYTL--SEGSNPEDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTESS 617

Db 603 ysyfmrllalepslkrkfdvpglrrk--eivdkipyyspnsitp-tpkktdds 654

Search completed: June 13, 2002, 08:42:39
Job time: 125 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:41:44 ; Search time 13.62 seconds
(without alignments)
1893.335 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484

Sequence: 1 KIYDNKNQLIADLGSRVYN.....TQSNTPDQONPOPAQP 666

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3484	100.0	719	1	BPBA_STRPN
2	2698.5	77.5	637	1	BPBA_STROR
3	999	28.7	914	1	BPBA_BACSU
4	739	21.2	714	1	BPBF_BACSU
5	712	20.4	726	1	BPBA_AQUAE
6	680.5	19.5	798	1	BPBA_NEIFL
7	642.5	18.4	822	1	BPBA_PSEAE
8	640	18.4	798	1	BPBA_NEIFA
9	640	18.4	798	1	BPBA_NEIMA
10	635.5	18.2	797	1	BPBA_NEIGO
11	633	18.2	798	1	BPBA_NEICI
12	617	17.7	825	1	BPBA_VIBCH
13	598.5	17.2	624	1	BPBD_BACSU
14	591.5	17.0	853	1	BPBA_HAEIN
15	584	16.8	777	1	BPBP_VIBCH
16	567	16.3	850	1	BPBA_ECOLI
17	526	15.1	844	1	BPBP_ECOLI
18	525.5	15.1	781	1	BPBP_HAEIN
19	502.5	14.4	787	1	BPBA_RICPR
20	490.5	14.1	809	1	BPBA_XYLFA
21	472	13.5	760	1	BPBP_BUCAI
22	450.5	12.9	770	1	BPBC_ECOLI
23	312.5	9.0	207	1	TRG_ALCEU
24	207	5.9	716	1	BPBP_BACSU
25	206	5.9	233	1	MTGA_NEIGO
26	206	5.9	645	1	SP5D_BACSU
27	194.5	5.6	224	1	MTGA_ACICA
28	193.5	5.6	598	1	FTSI_MESVI
29	192	5.5	233	1	MTGA_NEIMA
30	185.5	5.3	242	1	MTGA_KLEPN
31	183.5	5.3	242	1	MTGA_ECOLI
32	181	5.2	246	1	MTGA_HAEIN
33	164	4.7	588	1	FTSI_ECOLI

RESULT 1

ID	BPBA_STRPN	STANDARD:	PRT:	719 AA.
AC	Q04707			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Penicillin-binding protein 1A (BPB-1A) (Exported protein 2).			
GN	PONA OR EXP2 OR SP0369.			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID:1313;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=45607, AND 63915;			
RX	MEDLINE=93010977; PubMed=1396576;			
RA	Martin C., Sibold C., Hakenbeck R.;			
RT	"Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain."			
RL	EMBO J. 11:3831-3836(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=R6;			
RX	MEDLINE=92325042; PubMed=1624444;			
RA	Martin C., Briese T., Hakenbeck R.;			
RT	"Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus pneumoniae and Streptococcus oralis with high homology to Escherichia coli penicillin-binding proteins 1a and 1b."			
RL	J. Bacteriol. 174:4517-4523(1992).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=TIGR4;			
RX	MEDLINE=21357209; PubMed=11463916;			
RA	Tetelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D., Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J., Durkin A.S., White M., Kolonay J.F., Nelson W.C., Peterson J.D., Umayam L.A., Whitt O., Salzberg S.L., Lewis M.R., Radune D., Holtzaple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L., McDonald L.A., Feldblyum T.V., Angiuoli S., Dickson T., Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C., Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;			
RT	"Complete genome sequence of a virulent isolate of Streptococcus pneumoniae."			
RL	Science 293:498-506(2001).			
RN	[4]			
RP	SEQUENCE OF 293-369 FROM N.A.			
RC	STRAIN=R6X;			
RX	MEDLINE=95020625; PubMed=7934910;			
RA	Pearce B.J., Yin Y.B., Masure H.R.;			
RT	"Genetic identification of exported proteins in Streptococcus pneumoniae."			
RL	Mol. Microbiol. 9:1037-1050(1993).			
CC	!- FUNCTION: CELL WALL FORMATION.			
CC	!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.			

34	147	4.2	1861	1	APU_THETU	P38536 t amylopull
35	142.5	4.1	750	1	BPBX_STRPN	P14677 streptococc
36	142	4.1	610	1	FTSI_HAEIN	P45059 haemophilus
37	141	4.0	716	1	YQGF_BACSU	P54488 bacillus su
38	141	4.0	2843	1	APC_HUMAN	P25054 homo sapien
39	140.5	4.0	516	1	PLB2_YEAST	P13692 enterococcu
40	139.5	4.0	706	1	DEXT_STRDO	Q03674 saccharomyc
41	138	4.0	1337	1	CH13_CANAL	P39653 streptococc
42	136.5	3.9	567	1	FTSI_BUCAI	P40954 candida alb
43	136	3.9	579	1	FTSI_BUCAI	P57317 buchnera ap
44	135.5	3.9	1007	1	Y741_CHLMU	O9pj16 chlamydia m
45	135	3.9	1365	1	GTFS_STRDO	P29336 streptococc

ALIGNMENTS


```
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: THESE STRAINS ARE PENICILLIN-SENSITIVE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X67873; CAA48073.1; -
DR EMBL: X67872; CAA48072.1; -
DR EMBL: M90527; AAA26956.1; -
DR EMBL: AE007349; AAK74536.1; -
DR TIGR: S28038; S28038.
DR TIGR: SP0369; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
KW Multifunctional enzyme; Complete proteome.
FT ACT_SITE 370 370 ACYLATED BY PENICILLIN (BY SIMILARITY).
FT DOMAIN 658 683 SER-RICH.
FT VARIANT 124 124 T -> A (IN STRAIN R6).
FT VARIANT 386 386 V -> I (IN STRAIN 63915).
FT VARIANT 388 388 D -> E (IN STRAIN R6).
FT VARIANT 397 397 E -> K (IN STRAIN 63915).
FT VARIANT 523 523 M -> I (IN STRAIN 63915).
FT VARIANT 533 533 D -> E (IN STRAIN 45607).
FT VARIANT 540 540 T -> S (IN STRAINS 63915 AND R6).
FT VARIANT 657 657 N -> S (IN STRAINS 45607 AND R6).
SQ SEQUENCE 719 AA; 79758 MW; 59D397E83B4B3AA6 CRC64;

Query Match 100.0%; Score 3484; DB 1; Length 719;
Best Local Similarity 100.0%; Pred. No. 2.5e-193;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFFDHRGIDTIRILGAFURN 60
DB 54 KIYDNKQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFFDHRGIDTIRILGAFURN 113
QY 61 LQSNLSGGSTLTQOLIKLTFTSTQDTISRKAQEAWLAIQLEQKATQOEILTYINK 120
DB 114 LQSNLSGGSTLTQOLIKLTFTSTQDTISRKAQEAWLAIQLEQKATQOEILTYINK 173
QY 121 VYMSNGYGMQTAQNYGKDLNLSLPQLALLAGMPQAPNQDPYSHPEAAQDRRLVL 180
DB 174 VYMSNGYGMQTAQNYGKDLNLSLPQLALLAGMPQAPNQDPYSHPEAAQDRRLVL 233
QY 181 SEMNKGVIISAEQYKAVNTPTDGLQSLKSASNPAYMDNLYKEVINQVEEETCYNLLT 240
DB 234 SEMNKGVIISAEQYKAVNTPTDGLQSLKSASNPAYMDNLYKEVINQVEEETCYNLLT 293
QY 241 TGMVYTNVDQEAQKHLMDIYNTDEYVAYPDELOVASTIYDVSNKGVIAGLGRHOSNN 300
DB 294 TGMVYTNVDQEAQKHLMDIYNTDEYVAYPDELOVASTIYDVSNKGVIAGLGRHOSNN 353
QY 301 VSFGINQAVETNRDWSGTMKPTIDYAPALEYGVYDSTATIVHDEPYNYPGNTPTVYNWDR 360
DB 354 VSFGINQAVETNRDWSGTMKPTIDYAPALEYGVYDSTATIVHDEPYNYPGNTPTVYNWDR 413
QY 361 GYFGNITLQYALQOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHSYNAISSNTTESD 420
DB 414 GYFGNITLQYALQOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHSYNAISSNTTESD 473
QY 421 KKYGASSEKMAAAYAAAFANGGTYTKPMYIHKVVFSDGSEKFEFNVGTRAMKETTAYMMTD 480
DB 474 KKYGASSEKMAAAYAAAFANGGTYTKPMYIHKVVFSDGSEKFEFNVGTRAMKETTAYMMTD 533
```

```
QY 481 MMKTVLTGTCRNAYLAWLPQAGKTGTSNYTDEELENHIKTSQFVAPDELFAGYTRKYSM 540
DB 534 MMKTVLTGTCRNAYLAWLPQAGKTGTSNYTDEELENHIKTSQFVAPDELFAGYTRKYSM 593
QY 541 AVWTGYSNRLTPLVGLTVAAKVYRSMTYLSGSGNPEDWNIPEGLYRNGEEFVKNGAR 600
DB 594 AVWTGYSNRLTPLVGLTVAAKVYRSMTYLSGSGNPEDWNIPEGLYRNGEEFVKNGAR 653
QY 601 STWNSPAPQPPSTESSSSSSSDTSSTSSSTPTPTNNSTTTNNNTTQOSNTTDPQQNQ 660
DB 654 STWNSPAPQPPSTESSSSSSSDTSSTSSSTPTPTNNSTTTNNNTTQOSNTTDPQQNQ 713
QY 661 PQAQP 666
DB 714 PQAQP 719

RESULT 2
PBPA_STOR ID PBPA_STOR STANDARD; PRT; 637 AA.
AC Q00573;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Penicillin-binding protein 1A (PBP-1A) (Fragment).
GN PONA.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92325042; PubMed=1624444;
RA Martin C., Briesse T., Hakenbeck R.;
RT "Nucleotide sequences of genes encoding penicillin-binding proteins
RT from Streptococcus pneumoniae and Streptococcus oralis with high
RT homology to Escherichia coli penicillin-binding proteins 1a and 1b.";
RL J. Bacteriol. 174:4517-4523(1992).
CC -!- FUNCTION: CELL WALL FORMATION.
CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M90528; AAA26958.1; -
DR PIR: B42893; B42893.
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
KW Multifunctional enzyme.
FT ACT_SITE 371 371 ACYLATED BY PENICILLIN (BY SIMILARITY).
FT NON_TER 637 637
SQ SEQUENCE 637 AA; 70891 MW; A6D198BCEA603A63 CRC64;

Query Match 77.5%; Score 2698.5; DB 1; Length 637;
Best Local Similarity 87.5%; Pred. No. 3.5e-148;
Matches 511; Conservative 39; Mismatches 33; Indels 1; Gaps 1;

QY 1 KIYDNKQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFFDHRGIDTIRILGAFURN 60
DB 54 KIYDNKQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFFDHRGIDTIRILGAFURN 113
QY 61 LQ-SNSLOGGSTLTQOLIKLTFTSTQDTISRKAQEAWLAIQLEQKATQOEILTYIN 119
```



```
Db 114 LRGGGLQAGASTLTQOLIKLYFTSTSDQTLRSRAQEAWLAVQLQEQATKOEILTYVIN 173
Qy 120 KYVMSNGYGMOTAAQNYVYGDNLNLSIPQALLAGMPOAPNQYDPYSHPEAAQDRNLV 179
Db 174 KYVMSNGYGMOTAAQNYVYGDNLNLSIPQALLAGMPOAPNQYDPYSHPEAAQDRNLV 233
Qy 180 LSEKNGOYISAEQYEKAVNPTITDGLQSLKASNPAYMDNLYKEVINQVEEETGYNLL 239
Db 234 LSEKNGOYITAEQYEKAINPTITDGLQSLKASNPAYMDNLYKEVIDQVEQETGYNLL 293
Qy 240 TTGMVYVNDQEAQKHLWDIYNTDEYVAYPDDELQVASTVDVNSGVKVIAGLGARHQS 299
Db 294 TTGMVYVNDVSKVOORLWDIYNTDEYVAYPDDELQVASTVDVNSGVKVIAGLGARHQS 353
Qy 300 NVSFGINQAVETNRDWSGTMKPTIDYAPALEYGVVDSTATVHDEPPYNYPGTNTVYVND 359
Db 354 NVSFGINQAVETNRDWSGTMKPTIDYAPALEYGVVDSTATVHDEPPYNYPGTNTVYVND 413
Qy 360 RGYFGNITLQYALQOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHYNAISSNTTSES 419
Db 414 KSYFGNITLQYALQOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHYNAISSNTTSES 473
Qy 420 DKKGASSEKMAAAYAAAFANGTYTKPMYIKHVFSDSGEKEFNSVNGTRAMKETAYVMT 479
Db 474 DKKGASSEKMAAAYAAAFANGTYTKPMYIKHVFSDSGEKEFNSVNGTRAMKETAYVMT 533
Qy 480 DMKKTLYTGYGRNAYLAWLPQAGKTGTSNVTDEIENHIKTSOFVAPDELPAQVTRKYS 539
Db 534 DMKKTLYTGYGRNAYLAWLPQAGKTGTSNVTDEIENHIKTSOFVAPDELPAQVTRKYS 593
Qy 540 MAVWTGYSNRLTPLVNGTLTVAAKYVRSMTYLSGNSPEDWNI 583
Db 594 MAVWTGYSNRLTPIVGDGYFAAKYVRSMTYLSGNSPEDWNI 637

RESULT 3
PBPA_BACSU STANDARD; PRT; 914 AA.
AC P39793:
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Penicillin-binding protein 1A/1B (PBP1) [Includes: Penicillin-
DE insensitive transglycosylase (EC 2.4.2.-) (peptidoglycan TGase);
DE penicillin-sensitive transpeptidase (EC 3.4.-.-) (DD-transpeptidase)].
GN PONA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 499-515.
RC STRAIN=168;
RX MEDLINE=95113769; PubMed=7814321;
RA Popham D.L., Setlow P.;
RT "Cloning, nucleotide sequence, and mutagenesis of the Bacillus
RT subtilis ponA operon, which codes for penicillin-binding protein
RT (PBP) 1 and a PBP-related factor.";
RL J. Bacteriol. 177:326-335(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / MARRBURG;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serrif P.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
RT the serA and kdg loci cloned in a yeast artificial chromosome.";
RL Microbiology 142:2005-2016(1996).
RN [3]
RP GROWTH REQUIREMENTS.
RC STRAIN=168;
RX MEDLINE=98389671; PubMed=9721295;
```

```
RA Murray T., Popham D.L., Setlow P.;
RT "Bacillus subtilis cells lacking penicillin-binding protein 1 require
RT increased levels of divalent cations for growth.";
RL J. Bacteriol. 180:4555-4563(1998).
RN [4]
RP SUBCELLULAR LOCATION.
RC STRAIN=168;
RX MEDLINE=99255546; PubMed=10322023;
RA Pedersen L.B., Angert E.R., Setlow P.;
RT "Septal localization of penicillin-binding protein 1 in Bacillus
RT subtilis.";
RL J. Bacteriol. 181:3201-3211(1999).
CC -!- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
CC TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
CC SUBUNITS) (BY SIMILARITY).
CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN, PROBABLY FOUND ALL
CC OVER THE WHOLE CELL AT LOW CONCENTRATIONS. ALSO LOCALIZES TO THE
CC DIVISION SITE IN VEGETATIVE CELLS.
CC -!- DEVELOPMENTAL STAGE: EXPRESSION IS CONSTANT DURING GROWTH,
CC DECREASES DURING SPORULATION AND IS INDUCED APPROXIMATELY 15 MIN
CC INTO SPORE GERMINATION.
CC -!- PTM: THE PRODUCT EXPRESSED FROM THE TRANSLATION OF THE PONA GENE
CC APPEARS AS TWO BANDS ON A GEL (1A AND 1b), BUT THE SPECIFIC AMINO
CC ACID SEQUENCE OF EACH PROTEIN IS UNKNOWN.
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC -!- MISCELLANEOUS: CELLS LACKING THE PROTEIN REQUIRE INCREASED LEVELS
CC OF MG2+ OR CA2+ FOR GROWTH AND GERMINATION. APPROXIMATELY 50% OF
CC CELLS WITHOUT THE PROTEIN CONTAIN ABNORMAL FTSZ RINGS, SUGGESTING
CC IT IS INVOLVED IN SEPTUM SYNTHESIS; INCREASED LEVELS OF MG2+ OR
CC CA2+ ONLY PARTIALLY ELIMINATE THE SEPTATION DEFECTS.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC TRANSGLYCOSYLASE FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC TRANSEPTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U11883; AAA64947.1; -
CC EMBL; L47838; AAB38459.1; -
CC EMBL; Z99115; CAB14148.1; -
CC HSP; P02751; 1FNA.
CC Subtilist; BG10954; pona.
CC InterPro; IPR003961; FN.III.
CC InterPro; IPR001264; Transglycosyl.
CC InterPro; IPR001460; Transpeptidase.
CC Pfam; PF00041; fn3; 1.
CC Pfam; PF00912; Transglycosyl; 1.
CC Pfam; PF00905; Transpeptidase; 1.
CC ProDom; PD001895; Transglycosyl; 1.
CC SMART; SM00060; FN3; 1.
CC Peptidoglycan synthesis; Cell wall; Transferase; Glycosyltransferase;
CC Hydrolase; Multifunctional enzyme; Transmembrane; Signal-anchor;
CC Antibiotic resistance; Complete proteome.
CC SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
CC POTENTIAL.
CC DOMAIN 1 37
CC TRANSMEM 38 58
CC DOMAIN 59 914
CC EXTRACELLULAR (POTENTIAL).
CC TRANSGLYCOSYLASE.
CC TRANSPEPTIDASE.
CC ACT_SITE 329 662
CC ACYLATED BY PENICILLIN (BY SIMILARITY).
CC SEQUENCE 914 AA; 99562 MW; 6978E33DFE2423B6 CRC64;
CC SQ
```


Query Match	28.7%	Score 999;	DB 1;	Length 914;
Best Local Similarity	31.2%	Pred. No. 4.2e-50;		
Matches 254;	Conservative 131;	Mismatches 256;	Indels 174;	Gaps 24;
Qy	2	IYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREDFHRGIDTILRGAFLNL	61	
Db	79	IYDNKNGEIAEVAEKRTYVSEIDPDVVVKEAFATATEDARFYEHHGIDPDVRIGGALVANF	138	
Qy	62	QSN-SLOGGSLTLOOLIKLTFTSTDSQTSRKAQEAWLAIQLEOKATQOEILTYINK	120	
Db	139	KDGFAGEGSITQOVVANSLLS---HOKTLRKVKQVQEWLSIQLEARNYSKDEILEMYLNR	195	
Qy	121	YVMSNGNYGMOATAQNYG-KDLNNLSLPQALLAGMPQAPNOYDPYSGHPAAQDRRLV	179	
Db	196	IYFSPRAYIGKAAEEFPGVTDLSKLTVEQAATLAGMPQSTAYNPVKNPDKAEKRRNV	255	
Qy	180	LSEKNQGYISAEQYKAVNPTID-GLQSLK-----SASNPAYMDNLYKEVINQVEE	233	
Db	256	LSLMKQGFISDSQYNKAKKAVKDEGVVYSQYKESASTNKYSAPVE---EVMKEIDEK	311	
Qy	234	TGYNLLTGMVDYTNVDOEAQKHLWDIYNTDEYVAYPDDELQVASTIVDSNGKVIQAOLG	293	
Db	312	SDVPSADGLKIYTLDTKQAKDLDELMDGOT-VGFTEG-MQGGVTLTLDTKNGEVRALGA	369	
Qy	294	ARHOSNVSGFNGAVETNRDWSMTKPTIDYAPALEYGVYDSTATIVHDPEYNYPGTNT	353	
Db	370	GRNOPVG---GFNVATQPKAOPGSIKPKILDYGPVIEKNK-STEYEQDDSDAYTY-SNGK	424	
Qy	354	PVYNWRCYFGNITLOYLALQOSRNPVAVETLNKVLGNLAKTFLNGLGIDYPSIHSNAIS	413	
Db	425	PIRDMDRYLGPISMRVYALASRNIPALKAFOAVGKDTAVDFANGLGLTKDNVTEAYS	484	
Qy	414	SNTTESDKYKASSSEKMAAAFAAFANGTYKPKMYIKHWFSDGSEKFSNVGTRAMKET	473	
Db	485	IGGFGEND--GVSPITMAGAYSACNNGCTYNEPHFKVSIENNDGTLKLTTPKSKSAMS	542	
Qy	474	TAYMTDMKTVLYGTGRNAYLAWLPQAGKTGTSNYTDEIEHNKTSQVAPDELPAQ	533	
Db	543	TAFMITDLKTAVTGTLQALQVQVEVAGKTGTTFDNEVKRY-NIASGGARDSWFGV	601	
Qy	534	YTRYSMAVWTVGYSNR---LPLVNGULTVAKVYRSMTYLSGCS---NPEDWNIEG	586	
Db	602	YTPQITAAVWTVGMGENEAGKSLSAEOKVAKRIFAQLIADVDGSGFERPD--SVVEA	659	
Qy	587	LYRNG-----BEVFKNGARST-----	602	
Db	660	TVEKSNPAKLAPNTPSKDKLTFYFKGTAPSTVSKTYEKEEETAKSLGNVYDKD	719	
Qy	603	-----W-----NSPAQ-----QPPS-----	613	
Db	720	NQSLTSLWNYGDATFAVKQSVGSGSYSEIQNSSAKEAVISGVOPGYSVKYKFEVTA	779	
Qy	614	-----TESSSSSSDSSTSQSSSTTPTST	635	
Db	780	KSTASTSEYVPAEKEDDEKDKQOQDDEKQDEKTDQDDTQDDSQKDGQDQDDTDDST	839	
Qy	636	N-----NSTTNP--NNNTQOSTTTPDQONQNPQ	662	
Db	840	NQDQKQDQNTNTPSDNNNQDSNDN-DNDNSNQ	873	
RESULT	4			
BPPE_BACSU		STANDARD;	PRT;	714 AA.
ID	BPPE_BACSU			
AC	P38050;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Penicillin-binding protein 1F (PBP-1F).			
GN	PBP OR PONA.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
OC	Bacillus/Staphylococcus group; Bacillus.			

OX	NCBI_TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
EX	MEDLINE=98240224; PubMed=9579061;			
RA	Noback M.A., Holsappel S., Kiewiet R., Terpstra P., Wambutt R.,			
RA	Wedler H., Venema G., Bron S.;			
RT	"The 172 kb prkA-addAB region from 83 degrees to 97 degrees of the			
RT	Bacillus subtilis chromosome contains several dysfunctional genes,			
RT	the glyB marker, many genes encoding transporter proteins, and the			
RT	ubiquitous hit gene.";			
RL	Microbiology 144:859-875(1998).			
[2]				
RN	SEQUENCE OF 1-129 FROM N.A.			
RP	STRAIN=168;			
RC	MEDLINE=93328693; PubMed=8335642;			
EX	Popham D.L., Setlow P.;			
RA	"Cloning, nucleotide sequence, and regulation of the Bacillus			
RT	subtilis pbpF gene, which codes for a putative class A			
RT	high-molecular-weight penicillin-binding protein.";			
RL	J. Bacteriol. 175:4870-4876(1993).			
[3]				
RN	SEQUENCE OF 122-714 FROM N.A.			
RP	MEDLINE=93094140; PubMed=1459957;			
RA	Hansson M., Hedersstedt L.;			
RT	"Cloning and characterization of the Bacillus subtilis hemery gene			
RT	cluster, which encodes protoheme IX biosynthetic enzymes.";			
RL	J. Bacteriol. 174:8081-8093(1992).			
CC	-!- FUNCTION: CELL WALL FORMATION. MAY BE INVOLVED IN OUTGROWTH OF THE			
CC	GERMINATED SPORE OR IT COULD FUNCTION IN THE SYNTHESIS OF THE GERM			
CC	CELL WALL.			
CC	-!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.			
CC	-!- SUBCELLULAR LOCATION: Membrane-associated.			
CC	-!- DEVELOPMENTAL STAGE: EXPRESSION REMAINS CONSTANT DURING VEGETATIVE			
CC	GROWTH, DECREASES DURING EARLY SPOULATION, AND IS INDUCED IN THE			
CC	FORESPORE DURING LATE SPOULATION.			
CC	-!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE			
CC	TRANSGLYCOSYLASE FAMILY.			
CC	-!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE			
CC	PROTEINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; Y14083; CAA74517.1; -			
DR	EMBL; L10630; AAA71942.1; -			
DR	EMBL; M97208; AAA22516.1; -			
DR	EMBL; Z99109; CAB12851.1; -			
DR	PIR; A40614; A40614.			
DR	Subtilisin; BG10428; pbpF.			
DR	InterPro; IPR001264; Transglycosyl.			
DR	InterPro; IPR001460; Transpeptidase.			
DR	Pfam; PF00912; Transglycosyl; 1.			
DR	Pfam; PF00905; Transpeptidase; 1.			
DR	ProDom; PD001895; transglycosyl; 1.			
DR	Peptidoglycan synthesis; Cell wall; Transmembrane; Signal-anchor;			
KW	Complete proteome.			
FT	DOMAIN 1 12			
FT	TRANSMEM 13 33			
FT	POTENTIAL..			
FT	POTENTIAL..			
FT	EXTRACELLULAR (POTENTIAL).			
FT	DOMAIN 34 714			
FT	DOMAIN 49 217			
FT	TRANSGLYCOSYLASE.			
FT	TRANSPEPTIDASE.			
FT	ACT SITE 297 592			
FT	ACYLATED BY PENICILLIN (BY SIMILARITY).			
FT	SEQUENCE 714 AA; 79278 MW; 08D96718C84BB434 CRC64;			
SO				


```
Query Match      21.2%; Score 739; DB 1; Length 714;
Best Local Similarity 32.2%; Pred. No. 2.7e-35;
Matches 187; Conservative 119; Mismatches 37; Indels 38; Gaps 18;

QY 1 KIYDNKQLADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFARN 60
DB 50 KIVDQNGDEVASLYTENREPVSIENIEIPQVREAFIAVEDKRFYEHGIDAKSVGRVYRD 109
QY 61 -LQSNLSQGGSTLQOLIKLTYFSTSDQTSRKAQFAWLAIOLEOKATKOEILTYIN 119
DB 110 ILAGGKVEGGTTITQOLAKNLF--THDKTLRKTEVIITAINLERDYSKDKLLEMYLN 166
QY 120 KVMNSNGYGMQTAQNYGKDLNLSLPQALLAGMPQAPNQDYPDPSHPAAQDRNLV 179
DB 167 QLYFGHGVYGIQAASHYFNKFEVDLTVSEGAVALAIPKAPSTYSPILHPDKNERRDTI 226
QY 180 LSEMNKNGYISAEQYKAVNTPIPTDGLQSLKASNYPAYMDNLYLKEVINQVEEB--TGY 236
DB 227 LGMNDQGYISAKAVTAQGR--TLGLH-VKKQSETP-WFDSYIDLVIKEAEDKYSISGE 282
QY 237 NLFTTGMVDVTVNDOEAQKHLWDINTDEYVAYPDDELQVASTIVDVSNQKVIQAQLGARH 296
DB 283 QLLOGGYTIKVPDLKQKLTAYQVWKESGYYPGTDQNAEGSAVFINKTKTGGVEAIGRD 342
QY 297 QSSNVSGINQAVETNRDGMSTKPIITDYAPALEYGVYDSTATIVHDEPYNYPGTNTFVY 356
DB 343 YTSK--GYNR-VTAVRQPGSTFKPLAVYGPAMQKKF-KPYSLLKDLQSY-GDYTP-K 395
QY 357 NWDRGYGNITLQALQOSNRVPAVEILNKLGNRAKTFELNGLGIDIPSIHYNAISSNT 416
DB 396 NYDSRYEGEVMTSDAITYSKNAPAVTLNEIGVETGSKYLRANGIDIPDEGLALALGG-- 453
QY 417 TESDKKYGASKEAAVAAFAANGTGYKPMYIHKVVFSDGSE--KEFSNVCTRAMKETTA 475
DB 454 ---LEKGVSPQLAGAFHFAANGTTEPFPISSIDEDGETTADHKEKRVFSKOTS 509
QY 476 YMMTDMKTVLTYTGTRN-AYLAWLPQAGKTGTSNTYDEEIHKTSSQFVAPDELFAGY 534
DB 510 WNMTRLMQVVKKGATSGTYHGDL--AGKTGSTSYTG--VSGATK-----DAWPAFY 558
QY 535 TRKYSMAVWTGYSNR-LTPLYNGLITVAAKVYRSMYILSE 574
DB 559 TPKITGAVMMGYDKTDQNHLYLKAGSSYPTRLFKDILTQAGE 599

RESULT 5
BPBA_AQUAE STANDARD; PRT; 726 AA.
AC 066874;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Penicillin-binding protein 1A (PBP-1a) (PBP1a) [Includes: Penicillin-
DE insensitive transglycosylase (EC 2.4.2.-) (peptidoglycan Tcase);
DE Penicillin-sensitive transpeptidase (EC 3.4.-.-) (DD-transpeptidase)].
GN MRCA OR PONA OR AQ_624.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Kellner M., Aubay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "the complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
```


Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
[2]

```

Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
[2]
RN      SEQUENCE FROM N.A.
RP      STRAIN=ATCC 15692 / PA01;
RX      MEDLINE=20437337; PubMed10984043;
RA      Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA      Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA      Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA      Brody L.B., Coulter S.H., Folger K.K., Kas A., Larbig K., Lim R.M.,
RA      Smith K.A., Spencer D.N., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA      Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT      "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT      opportunistic pathogen.";
RL      Nature 406:959-964(2000).
[3]
RN      SEQUENCE OF 1-167 FROM N.A.
RP      STRAIN=ATCC 15692 / PA01;
RX      MEDLINE=940491325; PubMed7901733;
RA      Martin P.R., Hobbs M., Free P.D., Jeske Y., Mattick J.S.;
RT      "Characterization of pilQ, a new gene required for the biogenesis of
RT      type 4 fimbriae in Pseudomonas aeruginosa.";
RL      Mol. Microbiol. 9:857-868(1993).
CC      1-1- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
CC      PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
CC      PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
CC      (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
CC      TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
CC      SUBUNITS) (BY SIMILARITY).
CC      1-1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC      1-1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By
CC      similarity).
CC      1-1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC      TRANSGLYCOSYLASE FAMILY.
CC      1-1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC      TRANSEPTIDASE FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/
CC      or send an email to license@isb-sib.ch).

```

	--	EMBL; U73780; AAB39710.1;	-	
DR DR		EMBL; AE004517; AAC08430.1;	-	
DR DR		EMBL; L13867; AAC36826.1;	-	
DR DR		InterPro; IPR001264; Transglycosyl.		
DR DR		InterPro; IPR001460; Transpeptidase.		
DR DR		Pfam; PF00912; Transglycosyl; 1.		
DR DR		Pfam; PF00905; Transpeptidase; 1.		
DR DR		ProDom; PD001895; Transglycosyl; 1.		
KW KW		Peptidoglycan synthetase; Cell wall;		
KW KW		Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;		
KW KW		Signal-anchor; Antibiotic resistance; Complete proteome.		
DOMAIN	1	CYTOPLASMIC (POTENTIAL).		
FT FT	5	SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)		
FT FT	6	(POTENTIAL).		
FT FT	27	PERIPLASMIC (POTENTIAL).		
FT FT	48	TRANSGLYCOSYLASE.		
FT FT	DOMAIN	403 744		
FT FT	ACT_SITE	461 461		
SQ SQ	SEQUENCE	822 AA; 911199 MW; OBIA3FGFA50492AA CRC64;		
		FOLDED BY PENICILLIN (BY SIMILARITY) .		

```

Query Match          18.4%; Score 642.5; DB 1; Length 822;
Best Local Similarity 25.4%; Pred. No. 1.4e-29;
Matches 205; Conservative 122; Mismatches 252; Indels 229; Gaps 25;

Qy 1 KIYDNKNLIADIGSERVNAQAINDTDLVKATVSIEDHFFDHRGIDTIRILGAFERN 60
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 49 KVSDEGKLITSEFGEMERTIRFADIPQDRIHALLSDEDDNFANHYGVGVKSLMRAAQL 108

```

QY	61	LQNSLIQ--GGSTLTQOLIKLTFSTSTSDOTISRKAAQAWLAIQLEQKATKQEBILTYIYN	119
Db	109	LKSGCHITQGGSTITMVAQ-NYF--LTNRSFRSRINELLALQTEROLATDEILELYVN	165
QY	120	KVTMSGNYGMOTAAQNYGKDLNNLSLPOLALLAGMPQAPNOVDPTSHPEAAQDRNLV	179
Db	166	KIYLGNNRAGIEAQAQVYIGKPIKOLUSIAEMAMTAGLPAKPSRNVPLNPTFRSTRERNWI	225
QY	180	LSEMKNGQYISAEQYKAYNTPTDGLQSLKSASNYPAYMDNYLKEVINQVEEET--GY	236
Db	226	LERMLKLGFDIQQRQYAAVEEPINASYHVQTELPNAPYIAEMARAEMVGRYGSAYTEGY	285
QY	237	NLLTTGMDVYTN-----VDQEAQ-----KHL-----	257
Db	286	KVITTVRSDLQNAASQSVRDGLIDYDQRHGYRGPETRLPGQTRDAWLKHYQQQRSIGGLE	345
QY	258	-----WD-----IYNTDEYVAYPDDELQV	276
Db	346	PAIVTQVEKSGIMVWTRDQKEAVTWSMKWARPFLSNNSMGPMRQPDADVAQAQDQIRV	405
QY	277	-----ASTIVDSNGKVIQAQLGA-RHQSSNVSYFGINQAVETNRDWS	317
Db	406	QRQEDGTLRFVQIPAAQSALISLDPKDGAIRSLVGFSFEQSN--YNRAIQAKRQPGS	461
QY	318	TMKPIPTYAPALEYGYVDSTATIVIDEPNY-----PGTNPVYNWDRGYFGN	365
Db	462	SFKPFI-YSAALDNGP--TAASLVNDAPITFVDEYLDKVMWRPKNDTNT-----FLGP	510
QY	366	ITLQYALQOSNRNPVAVETLNKVLGNRAKTLFNLGLIDYPSI--HYSNAISSNTTSDKKY	423
Db	511	IPUREALYKSRNMVSRVLQGLIGIERAISYITKFGFQDELPRNPFSLALGTAT-----	563
QY	424	GASSEKMAAAYAAAFANGGYYPKMTVHKVVFSDG-----SEKFEFN	464
Db	564	-VTPMEIAGNWSVFANGGKVNPPYIERIESRQGVQLVQANPPRPVVEEQVAADAEDAGN	622
QY	465	VG-----TRAMKETTAYMTDMKTLVITYGTGR	492
Db	623	PGDPEHPESAEGSGEIAQOVAAKAQTTPETTPAERIIDAARTAYIMTSLMDVIKRGTR	682
QY	493	NAY-LAWLPQAGKTGTSNYTDEIENHIKTSQFVAPDELFGYTRKYSMVAWTGYSNRLT	551
Db	683	RALALKRTDLAGKTGNTN-----DSKQGWFSGYNSDYVTSVMVGFDQPET	727
QY	552	PLVGN---GLTVAAKVYRSMNTYLSGSGSNPDWNTPEGLYRNGEYFKNGARSTWNSPAP	608
Db	728	--LGRREYGGTVALPTWIRYMGFALKDQPMHTMAEPPGI-----VSLRIDPVTGRSAAP	779
QY	609	QPPPTSSSSSSSDSTSSSTSTPSTN	636
Db	780	GTPCAIFYEMFKNED-----TPPSVN	799
RESULT 8			
PBPA_NEILA STANDARD; PRT: 798 AA.			
ID	PBPA_NEILA		
AC	O87579;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DE	Penicillin-binding protein 1A (PBP-1a) (PBP1a) [includes: Penicillin-		
DE	insensitive transglycosylase (EC 2.4.2.-) (peptidoglycan TGase);		
DE	penicillin-sensitive transpeptidase (EC 3.4.-) (DD-transpeptidase)].		
GN	MRCA OR PONA.		
OS	Neisseria lactamica.		
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
NCBI_TaxID=486;			
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NRL 3716;		
RA	Ropp P.A., Nicholas R.A.;		
RT	"Nucleotide sequence of the ponA gene encoding penicillin-binding		
RT	protein 1 from Neisseria lactamica.";		

RA Gill J., Scarlato V., Mesignani V., Pizza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
 RT MC58.";
 RL Science 287:1809-1815(2000).
 CC -!- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
 CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
 CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
 CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
 CC TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
 CC SUBUNITS) (BY SIMILARITY).
 CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By
 CC similarity).
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC TRANSGLYCOSYLASE FAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
 CC TRANSEPTIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U80933; AAB52541.1; -;
 DR EMBL; AL162753; CAB83943.1; -;
 DR EMBL; AF002530; AAF42144.1; -;
 DR TIGR; NWB1807; -;
 DR InterPro; IPR001264; Transglycosyl.
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF00912; Transglycosyl; 1.
 DR Pfam; PF00905; Transpeptidase; 1.
 DR ProDom; PD001895; Transglycosyl; 1.
 KW Peptidoglycan synthesis; Cell wall; Transferase; Glycosyltransferase;
 KW Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;
 KW Signal-anchor; Antibiotic resistance; Complete proteome.
 FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 10 30 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
 FT (POTENTIAL).
 FT DOMAIN 31 798 PERIPLASMIC (POTENTIAL).
 FT DOMAIN 50 218 TRANSGLYCOSYLASE.
 FT DOMAIN 378 700 TRANSEPTIDASE.
 FT ACT_SITE 461 461 ACYLATED BY PENICILLIN (BY SIMILARITY).
 SQ SEQUENCE 798 AA; 88147 MW; 0BAF4EF037A9977B CRC64;

Query Match 18.4%; Score 640; DB 1; Length 798;
 Best Local Similarity 26.1%; Pred. No. 1.6e-29;
 Matches 203; Conservative 121; Mismatches 272; Indels 182; Gaps 21;

QY 2 IYDNKNQLIADLGSRVNAQANDPTDLVKAIVSTEDHRFDHRCIDIRILGAFRLNL 61
 DB 52 IYDSAGEVIGMYGEORREFTKIGDFEVLNVAIVAAIDKRRFHWGVDVGMVARAVGN 111
 QY 62 QSNLSQ-GGSTLTQOLIKLTFTSTSDTISRKAQEAWLALQLEKATKQELLTYINK 120
 DB 112 VSGSVQSGASTITQVAKNFYLS---SEKTFTRKNEVLLAYKIEQSLSKILELYFNQ 168
 QY 121 VYMSGNGTGMQTAQNYGKDLNNLSLPOLLAGMPQAPNQYDPYSHPEAAQDRNLVL 180
 DB 169 IYLGQRAYGFAAQAQYFNKNVRDLTLAEAAAGLAPKAPSAYNPVNPRAKLRQKYL 228
 QY 181 SEMKNGQYISAEQYKAVNTPTDGLQSLKSANPAYMDN-----YLKEVINO-VEEE 233
 DB 229 NNMLEKMITVOORQALNEEL-----HVERFVRKIDQSALYAEVMVRQELYEK 277
 QY 234 TGYNLLTGMGVYTNVDQEAQK----- 255
 DB 278 YGEDAYTQGFKYTTVRADHQKVAATEALRNDRGSSYRGAENYIDLSKSEDEETV 337

QY 256 --HLWDIYNTEDEVV----- 267
 DB 338 SQYLSGLYTVDKMVPVLDVTKKKNVVIQLPGGRRVTLDRRALGFAARAVNNEKMGEDR 397
 QY 268 -----AYPDDELQVASTIVDYSNCKVIAQLCARHOSNVSGINQAVE 310
 DB 398 IRRGAVIRVKNNGRWAVVOEPLLQGLVSLDAKTGAVALVGVGYDFHST---FNRAVQ 454
 QY 311 TNRDWSGTMKPIITDYAPALEYGYVDSTATIVHDEPNYPCGN-----TPVYNWDRGYFG 364
 DB 455 AMRQPGSTKPFV-YSAALS-KGM--TASTVVDAPISLPKGPNGSVWTP-KNSDGRYSG 510
 QY 365 NITLOVALQOSRNPVAVETLNKVLNRAKTFNLGLIDYPSIIHYSNAISNTTESKYYG 424
 DB 511 YITLQALTASKNWSIRILMSIGVGAQYIRRFGRSSSELPASISALGTGET----- 565
 QY 425 ASSEKMAAAYAFANGTGYKPMYIHKVVFSDGSEK-----FSNVGTRAMKETAYMM 478
 DB 566 -TPLKVAEAYSFANGYRVSSHVVDKIDYDRGLRAQMLVAGONAPOAIDPRNAYIM 624
 QY 479 TDMKMTLVITYGGRN-AYLAWLPQAGKTGTSTNYTDEIEHNIKTSQFVAPDELFAGYTRK 537
 DB 625 YKIMQDVVRVGTARGAAALGRTDIAGKTGTN-----DN-----KDAWFGVGNPD 669
 QY 538 YSAVWVTGYSN-RLTPLVGNGLTAAKAVYKSMYTYLSEGSNPEDWNIPEGLY-RNGEYVF 595
 DB 670 VVTAVYIGDFKPKSMGRVGYGGTIAVPVVDYMRFAKLGKQKGMKMGECVGSNGEYFM 729
 QY 596 K-----NGARSTWNSPAQOPP-----STESSSSSDSSTSSSSSTPTNNST 639
 DB 730 KERWVTDPLGLDLSNGIAQPSPRAKEDDGGAAEGGQAADDEVQDMQETPVLPSNT 787

RESULT 10
 ID BBPA_NEIGO STANDARD; PRT; 797 AA.
 AC 005131;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Penicillin-binding protein 1A (PBP-1a) (BBP1a) [includes: Penicillin-
 DE insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan TGase);
 DE Penicillin-sensitive transpeptidase (EC 3.4.-) (DB-transpeptidase)].
 GN MRCA OR PONA.
 OS *Neisseria gonorrhoeae*.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; *Neisseria*.
 OX NCBI_TaxID=485;
 RP [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-25.
 RC STRAIN=FA19, AND ATCC 700825 / FA 1090;
 RX MEDLINE=97252514; PubMed=9098083;
 RA Ropp P.A., Nicholas R.A.;
 RT "Cloning and characterization of the ponA gene encoding penicillin-
 RT binding protein 1 from *Neisseria gonorrhoeae* and *Neisseria*
 RT *meningitidis*.";
 RL J. Bacteriol. 179:2783-2787(1997).
 CC -!- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
 CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
 CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
 CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
 CC TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
 CC SUBUNITS) (BY SIMILARITY). ESSENTIAL FOR CELL WALL SYNTHESIS.
 CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By
 CC similarity).
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC TRANSGLYCOSYLASE FAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
 CC TRANSEPTIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; 072876; AAB52536.1; -
 InterPro; IPR001264; Transglycosyl.
 InterPro; IPR001460; Transpeptidase.
 Pfam; PF00912; Transglycosyl; 1
 Pfam; PF00905; Transpeptidase; 1.
 ProDom; PD001895; Transglycosyl; 1.
 Peptidoglycan synthesis; Cell wall; Transferase; Glycosyltransferase;
 Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;
 Signal-anchor; Antibiotic resistance.
 INTP MET 0 0
 DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 9 29 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN) (POTENTIAL).
 DOMAIN 30 797 PERIPLASMIC (POTENTIAL).
 DOMAIN 49 217 TRANSGLYCOSYLASE.
 DOMAIN 377 699 TRANSEPTIDASE.
 ACT_SITE 460 460 ACYLATED BY PENICILLIN (BY SIMILARITY).
 SEQUENCE 797 AA; 88364 MW; C7A01D2B1CAC9F3B CRC64;
 Query Match 18.2%; Score 635.5; DB 1; Length 797;
 Best Local Similarity 26.4%; Pred. No. 2.9e-29;
 Matches 207; Conservative 120; Mismatches 287; Indels 169; Gaps 22;
 QY 2 IYDNKNQLIADGSRERRVNAQANDIPTDLVKAIVSIEDHRRFDHRTIRILGAFRLNL 61
 DB 51 IYSDAGEVIGMYGEORREFTKIGDFEVLNRNAVIAEDKRYRHGVDVWGAARAVGNV 110
 QY 62 QNSLQ-GGSTITQOLIKLTFTSTSDTISRKAQEWLAQLQEKATKQBIILYYINK 120
 DB 111 VSGVSGASTITQOVANKFYLS-----SEKTFTRKNEVLAYKIEQSLSKDILYFNQ 167
 QY 121 VYMSNGNTGMOTAPYQKGLNNLSLPQALLAGMPOAPNOYDPYSHPEAAQDRRLVL 180
 DB 168 IYLGORAYGFASAAQIYFNKNVRDLTAEAMLAGLPKAPSAYTNVPERAKLRQKYL 227
 QY 181 SEMKNGYISAEQYKAVNTPI--TDGLOSLSKASNYPAYMDNYLKEVINQVEETGYNL 238
 DB 228 NNMEEKMITVQORDQALNEELHYERFVRKIDQSALYAE-----VRRELYEKYGEDA 281
 QY 239 LTGMDVYTNVDQEAQ-----HLW 258
 DB 282 YTGFKVYTTVTRDHOKAATEALRKALNFRDGRSSYGAENYIDLKSEDVEETVSQYLS 341
 QY 259 DTYNTDEYV-----AYPDD----- 273
 DB 342 GLYTVDKMVPVAVLDVTKKNVVIQIPGRRVALDRALGFAARAVDNEKMGEDRIIRCA 401
 QY 274 -----LQVASTIVDSNGKVIQAGLARGHSSNVSGFQINQAVETNRDW 315
 DB 402 VIRKNGNGRWAVQOEPLQGLALVDARTKAVRALVGDFHSHKT-----FNRAVQAMRQP 458
 QY 316 GSTMKPITDYPALEYGVYDSTATVHDPEYNYPCFN-----TPVYNWRGTYFCNITLQ 369
 DB 459 GSTFFAFV-YSAALSCKM--TASTVVDNAPISLPKGGNGSVWTP-KNSDGRYSYITLR 514
 QY 370 VALQOSNRNPVAVETLKNKLNKRLFTGLNGLDIDYPSIHSNAISSNTESDKKYGASSEK 429
 DB 515 QALTASKNWSIRILMSIGVGVAQYQYIRFGFRPSELPSALSMALGTGT-----TPLK 568
 QY 430 MAAAYAAAFANGTYTKYKPIIHKVWFSDDSEK-----FSNVGTRAMKETTAAMTDMWK 483
 DB 569 VAEAYSFVANGGYRVSSSHVIDKIYDRDRLRAQMQLVAGQNAQIDPRNAYIMYKIMQ 628
 QY 484 TVLTGTGRN-AYLAWLQPAQKTGTSNYTDEIENHKTSQFVAPDELFAGYTRYKYSMAV 542
 DB 629 DVVRVGTARGAAALGRDITAGKTGTN-----DN-----KDAWFGVGNPDVVTAV 673

QY 543 WTGYSN-RLTPLVGNGLTVAAKVYRSMMTYLSGSSNPEDWNIPEGLY-RNGEYVFNK--- 597
 DB 674 YIGFDKPKSMRGAGTGTIAPVWVDYMRFAKLGKQKMGKPEGVVSSNGEYMKERWV 733
 QY 598 ---GARSTWNSAPQPPPTSSSSSSSTSSSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 654
 DB 734 TDPGLMLDNGIAPQ--PSRRAKEDDAEVAENEQQGRSDETRDQVQETP---VLPSTNIDS 788
 QY 655 DQO 657
 DB 789 KOQ 791
 RESULT 11
 PBPA_NEICI
 ID PBPA_NEICI STANDARD; PRT; 798 AA.
 AC 086088;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Penicillin-binding protein 1A (PBP-1a) (PBP1a) [Includes: Penicillin-
 DE insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan TGase);
 DE penicillin-sensitive transpeptidase (EC 3.4.-) (DD-transpeptidase)].
 GN MRCA OR PONA.
 OS Neisseria cinerea.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=483;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRL 30066;
 RA Ropp P.A., Nicholas R.A.;
 RT "Cloning and sequence analysis of the pona gene encoding penicillin
 RT binding protein 1 from Neisseria cinerea."
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
 CC PEPTIDOGYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
 CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
 CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
 CC TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
 CC SUBUNITS) (BY SIMILARITY).
 CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGYCAN SYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By
 CC similarity).
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC TRANSGLYCOSYLASE FAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
 CC TRANSEPTIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DB EMBL; AF085340; AAC34128.1;
 DR InterPro; IPR001264; Transglycosyl.
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF00912; Transglycosyl; 1.
 DR Pfam; PF00905; Transpeptidase; 1.
 DR ProDom; PD001895; Transglycosyl; 1.
 DR Peptidoglycan synthesis; Cell wall; Transferase; Glycosyltransferase;
 DR Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;
 DR Signal-anchor; Antibiotic resistance.
 FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 10 30 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN) (POTENTIAL).
 FT DOMAIN 31 798 PERIPLASMIC (POTENTIAL).
 FT DOMAIN 50 218 TRANSGLYCOSYLASE.
 FT DOMAIN 414 700 TRANSEPTIDASE.
 FT ACT_SITE 461 461 ACYLATED BY PENICILLIN (BY SIMILARITY).
 SQ SEQUENCE 798 AA; 87843 MW; 5842ED4BCB9FF06A CRC64;


```
Query Match      18.2%; Score 633; DB 1; Length 798;
Best Local Similarity 26.5%; Pred. No. 4.1e-29;
Matches 191; Conservative 115; Mismatches 248; Indels 166; Gaps 19;

QY 2 IYDNKNQIADLGSRRVNAQANDIPTDLVKAIVSIEDHRFFDHGIDTIRILGAFRLN 61
DB 52 IYSADEGEVIGIYGEORREFTKTGDFEVLNRNAVIAEDKRFYQHWGVDVWGVARAVVGN 111
QY 62 QNSLQ-QGSTLTQOLIKLTYFSTSDQTSRKAQEAWLAIQLEKATKQKQKQKQKQK 120
DB 112 VAGGVQSGASTITQVAKNFYLS---SEKTRKNEALLAYKIQSLSKQKQKQKQK 168
QY 121 VYMSNGNYGMOTAAQNYGKOLNLSLPQALLAGMPQAPNOYDPSHPPEAAQDRRLV 180
DB 169 IYLGORAYGFAAQIYFNKQVRELTLAEVAMLAGLPKAPSAYNPIVNPRAKQKQYL 228
QY 181 SEMKNGQYISAEQYKAVNTPITDGLQSLKSASNPAYMDN-----YLKEVINQ-VEE 233
DB 229 NMLEEKMITLQORQALNEEL-----HYERFVKIDQSALYVAEMVRQELYEK 277
QY 234 TCYNLLTTGMVYTVNDOBAQK----- 255
DB 278 YGEDAYTOGFKYTYTVRDHQVATEALRKALNFRDSSYRGARSYIDLKGEDVEETV 337
QY 256 --HLMDIYNTDE----- 265
DB 338 SOYLSGLYTVDRKWPALVDLTKRKNVYIQLPSGKRVTLDRSLGFAARAVNNEKMGESR 397
QY 266 -----YVAYPDDELQVASTVDVSGKVIQALGARHQSNSVSGINQAVE 310
DB 398 IRRGSVIRVRNNGRWVYVQEPPLQATLVSLDAKTCVAVRALVGYDFHSEKT---FNRAAQ 454
QY 311 TNRDNGSTMKPIITDAPALEGYVDYSTATIVHDEPNYP--CTNTPVY---NWRDGYFCN 365
DB 455 ANRQGSFTKPII-YSAALSKGM--TASTVNDADISLPKGKANGSVWTPKNSDGRYSY 511
QY 366 IFLOYALQOSRNPVAVETLNKVLGNRAKTLFNLGLGIDYPSIHYNSAISNTTESDKKYGA 425
DB 512 ITRQALTASKNWSIRILMSIGVGYAHEYIQRFQFKPSELPASLSMALGTGET----- 565
QY 426 SSEKAAAYAFANGSTYKPMYIHKVFSQSEKE-----FSNVGTRAMKETAYMMT 479
DB 566 TPLKTAEAYSFANGGYRVSSHVVDIKYSGDGLRQAQOPLVAGONAPOAIDPRNAYIMY 625
QY 480 DMKQVTLVYTGCRN-AYLAWLPQACKTGTSNYTDEIEHNHKTQFVAPDELFAGYTRYK 538
DB 626 KIMQDVVRVGTARGAALGRSDIAGKTGTN-----DN-----KDAWFGVFNPDV 670
QY 539 SMAVWTGYSN-RLTPLVNGLTVAARVSRMMTYLSEGSNPDWNIPEGLY-RNGEFVFK 596
DB 671 VTAVYIGFDKPKSMGRAGVGGTIAVPVWVDMRFALKGGQGMKVPGEVVSNGEYMK 730

RESULT 12
PBPA_VIBCH
ID PBPA_VIBCH STANDARD; PRT; 825 AA.
AC Q9KNU5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Penicillin-binding protein 1A (PBP-1a) (PBPA) [Includes: Penicillin-
DE insensitive transglycosylase (EC 2.4.2.-) (peptidoglycan TGase);
DE penicillin-sensitive transpeptidase (EC 3.4.-.-) (DD-transpeptidase)].
GN MRCA OR PONA OR VC2635.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
ON NCBI_TaxID=666;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
```

```
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Olin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -!- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
CC TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
CC SUBUNITS) (BY SIMILARITY).
CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By
CC similarity).
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC TRANSGLYCOSYLASE FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC TRANSEPTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE004330; AAF95776.1; ALT_INIT.
DR TIGR; VC2635.
DR InterPro; IPR001264; Transglycosyl.
DR Pfam; PF00912; Transglycosyl; 1.
DR Prodom; PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Cell wall; Transferase; Glycosyltransferase;
KW Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;
KW Signal-anchor; Antibiotic resistance; Complete proteome.
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 7 27 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT FT (POTENTIAL).
FT DOMAIN 28 825 PERIPLASMIC (POTENTIAL).
FT DOMAIN 48 216 TRANSGLYCOSYLASE.
FT DOMAIN 413 752 TRANSEPTIDASE.
FT ACT_SITE 471 471 ACYLATED BY PENICILLIN (BY SIMILARITY).
FT SEQUENCE 825 AA; 91986 MW; 8A9553DA2650B9D2 CRC64;
Query Match 17.7%; Score 617; DB 1; Length 825;
Best Local Similarity 24.3%; Pred. No. 3.6e-28;
Matches 200; Conservative 128; Mismatches 220; Indels 276; Gaps 28;

QY 1 KIYDNKNQIADLGSRRVNAQANDIPTDLVKAIVSIEDHRFFDHGIDTIRILGAFRLN 60
DB 49 QVFSQDGKLIQAQFGKRRPRLKLEEMPKELIEAVIATEDSRVYEHYGFDPGIGITRAFAFV 108
QY 61 LOSNSL-QGSTLTQOLIKLTYFSTSDQTSRKAQEAWLAIQLEKATKQKQKQKQK 119
DB 109 LASGASQSGASTITQOLARNFFLS---NEKVMYRKVEIFAIHTEQLLSKQKQKQK 165
QY 120 KYVMSNGNYGMOTAAQNYGKOLNLSLPQALLAGMPQAPNOYDPSHPPEAAQDRRLV 179
DB 166 KIYLGSRVGVGAQAQVFGKVKDLTLGEIALIAGLPKAPMTMNPPIYSEVATRNRVV 225
QY 180 LSEKNGQYISAEQYKAVNTPITDGLQSLKSASNPAYMDNLYKEV-----INQVEET 234
DB 226 LQRMLEDEKVIYTKAEYDAARAEPVLPVYHCAETELNAP-----YVAETARAWWVYGEFA 280
QY 235 GYNLLTTGMVYTVNDOE-----AQKHLWDI----- 260
DB 281 AY---TSGMNVYTVDSKLQRAANQAAINNLLAYDHRHGYRGAERKELMQVNPWSSQTL 337
```


THIS PAGE BLANK (USPTO)